DOCUMENT RESUME

ED 323 564 CS 212 522

AUTHOR	Smith, Carl B.; Ingersoll, Gary M.
TITLE	Written Vocabulary of Elementary School Pupils, Ages 6-14. Monograph in Language and Reading Studies Number 6.
INSTITUTION	Indiana Univ., Bloomington. School of Education.
REPORT NO	ISSN-0193-4759
PUB DATE	Jan 84
NOTE	50p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Basic Vocabulary; Elementary Education; *Elementary School Students; Reading Comprehension; *Word Lists; *Writing (Composition); Writing Evaluation; Writing Research
IDENTIFIERS	*Childrens Writing; Cultural Change; Spelling Errors

ABSTRACT

A study explored the written compositions of elementary school students (ages 6-14) and the vocabulary they use. Compositions were written by a large national sample of over 4,000 children, who were given free rein to write whatever they wanted; thus the study provides status information on the vocabulary that children currently use. The study also explored relationships between vocabulary and the communicativeness of the compositions and between vocabulary and reading comprehension. A major outcome of the study was the compilation of a lexicon of children's written vocabulary. That word list shows huge cultural changes in the vocabulary that children use when compared to word lists (still in common use today) which are 40 to 60 years old. Results also included analyses of token production, type production, type-token relationships, and spelling error production. (Two tables of data and six figures are included, and 107 references are attached. Two appendixes contain a list of the 500 most frequent words, and a list of common misspellings.) (SR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

monograph in language and reading studies
US ISSN 0193-4759 / Number 6 / January 1984

Written Vocabulary of Elementary School Pupils, Ages 6-14

Carl B. Smith, Ph.D.
Language Education Department

Gary M. Ingersoll, Ph.D.
Counseling and Educational Psychology Department

Monographs in Teaching and Learning / School of Education / Indiana University
US ISSN 0193-4740

ED323564

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Carl B. Smith
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

CS212528

Indiana University's Language Education Departments initiated the Monographs in Language and Reading Studies to serve a number of purposes, including the publication of papers based on conference proceedings, research and development reports, and papers of special interest to teachers and other professionals working in the language education areas.

Howard D. Mehlinger, Dean

David H. Gliessman, Editor: **Dorothy J. Taylor,** Associate Editor

Monograph in Language and Reading Studies

Series Editor

Leo C. Fay, Reading Education

Editorial Board

Robert C. Lafayette, Foreign Language Education; **Larry Mikulecky,** Language Education; **Karen Schuster-Webb,** Bilingual Education; **Sharon L. Smith,** Reading Education; **Vernon H. Smith,** English Education

Earlier monographs in this series may be obtained from Monographs in Teaching and Learning, School of Education, Indiana University, Bloomington, IN 47405. Make checks payable to Indiana University.

Secondary Reading; Theory and Application **(William Diehl et al., Eds.) $3.95**

Adult Reading Habits, Attitudes, and Motivations **(Larry Mikulecky) $3.00**

New Perspectives on Comprehension **(Jerome Harste and Robert Carey, Eds.) $6.95**

Reading English as a Second Language: Moving from Theory **(C. W. Twyford, William Diehl, and Karen Feathers, Eds.) $6.95**

Relating Reading and Writing: Developing a Transactional Theory of the Writing Process **(Nancy Leavitt Shanklin) $10.00**

Monographs in Teaching and Learning comprises series of monographs in a variety of disciplines. Each series is published on an irregular schedule by Indiana University School of Education. Articles are devoted to aspects of major current concern to educators in the respective fields and deal with theory, research, or practical application pertaining to all levels of practice.

Preface

The monograph series in Language Education was established to publish research and theoretical statements that cannot be accommodated in the limited space provided by the typical education journal. Previous monographs have treated secondary reading, adult reading habits, perspectives on comprehension, and reading English as a second language. This present monograph pushes into the little-explored area of the written compositions of elementary school children and the vocabulary children use.

This study by Smith and Ingersoll provides status information on the vocabulary that children use when they are given free rein to write whatever thay want. Relationships between vocabulary and the communicativeness of the compositions and with reading comprehension are also explored. A huge change has taken place in the written vocabulary of children over the past forty years. Smith and Ingersoll make quantitative comparisons between their list and those of Rinsland and others. In view of the stock placed in several word lists that are forty to sixty years old, curriculum developers and teachers can gain valuable insights for their work from the results reported in this study.

—Leo C. Fay

Introduction

In his assessment of writing research, Graves (1978, 1981) was forced to conclude that research and curricular development in children's writing has not been a major focus in American education. Efforts in language arts research and development are overwhelmingly aimed at reading. The wide disparity in effort is reflected in Graves's observation that for each $1 spent on teaching children to communicate through writing, schools spend nearly $3000 on teaching them to decode written communication—that is, to read. In one recent conference, Whitman (1981) went so far as to conclude that writing research lags 50 to 100 years behind reading research. The result of this disparity in effort has been a general paucity of knowledge about productive language behavior as seen in children's writing.

In many ways the lack of research in writing is curious since the link between children's productive vocabulary, language skills, and reading comprehension has been firmly established. One purpose of the research reported in this monograph is to respond to this void through a study of a large national sample of compositions written by children in grades 1 through 8, ages 6-14. Several products result from this study, not all of which are included herein. A major outcome of this research has been the compilation of a lexicon of children's written vocabulary. That word list, in turn, has been the subject of a series of quantitative and qualitative analyses. Additionally, the compositions themselves have been the subject of a variety of studies. The breadth of the sample of compositions available for analysis in this study is so wide that it will be several years before all research hypotheses are exhausted. For this report, attention is concentrated on the vocabulary production in these compositions and a limited set of relationships.

Composition Research

The vast majority of literature on composition deals with classroom techniques, lesson ideas, and motivation schemes. While these are interesting and useful for the practitioner, they do not address the critical issue of determining if, in fact, composition skills are being learned and refined over time.

For a historical perspective on the research in composition we recommend a report by VanDeWeghe (1978) in which he examines the research response to the Braddock Report (1963) which posed 24 research questions about written

composition. Another historical view is gained from Smith (1967) who outlined research and evaluation instruments dating from 1912. Those sources indicate a long-standing interest in holistic rating scales and, more recently, in the measurement of mechanistic aspects of composition, especially syntax.

Since research often depends on evaluation techniques, perhaps the relative dearth of research in composition can be explained by the two major criticisms of composition testing: lack of test instruments and inadequate and artificial test environment. In Buros' (1975) *English Tests and Reviews*, there are no normed overall composition tests targeted. Standardized tests deal instead with many discrete aspects of writing, such as grammar, punctuation, vocabulary, and spelling. While researchers agree that these factors certainly are highly correlated with the quality of writing, they do not, in and of themselves, measure what is and is not good composition.

Tests of writing mechanics as measures of quality of writing come under fire frequently in the literature. Newkirk (1977) lists five major criticisms of these writing tests:

—Many have little actual writing. They depend on multiple choice or fill-in responses.
—Most have questionable content validity. They do not establish that they really test what they purport to test.
—Most provide inadequate time for students to produce their best efforts.
—Most provide inadequate motivation for students to perform well.
—Most incorporate improper, irresponsible, or inaccurate methods of interpreting and using the results that are offered.

McCleary (1979) examines the topics of composition tests. Nowhere in the research does he find it demonstrated that the subjects about which the students are asked to write are reliable or valid. He hypothesizes that the student's experience with the topic, interest in the topic, and the manner in which the topic is introduced and explained will have profound influence upon the child's ability to perform.

Perron (1976) substantiates this view with a series of studies that show how the mode of discourse affects the quality of student writing as judged by syntactic maturity. He examined 52 students in each of the third, fourth, and fifth grades and found significant differences in their writing in the various modes: argumentation, exposition, narration, and description. Specifically he discovered that argumentation topics rendered the most complex and mature syntax and that high-ability students performed better on these modes than other students. Descriptive writing tended to be the least complex with exposition and narration in the middle. Perron cites many other studies that similarly found that an individual student will perform quite differently depending on the kind of writing task required. Lloyd-Jones (1977) also found that good writing in one mode does not necessarily mean equal success in another mode.

One should not, on the other hand, be overly ready to discard mechanistic variables in evaluating compositions. While such variables may lack the artistic appeal of more qualitative approaches, they provide a reliable base on which to make inference. Qualitative ratings, without a common base, run the risk of low reliability.

Spelling

Spelling skills among children continue to baffle educators. No one is sure exactly why some people seem to spell effortlessly and others resist all attempts to improve. In analyzing the mechanics of the spelling task, Simon (1976) noted that a two-part process is apparently at work. A child searches for a word in long-term memory and, failing to find it, generalizes a graphemic pattern from the options for that phoneme—i.e., sounds it out. She further classifies all spelling errors into four categories:

1. Errors of perception
   —from idiosyncratic mispronunciations, e.g., "warter" for "water"
   —homophones(e.g., "to" and "too")
   —unknown words which are attempted totally phonetically
2. Errors of generation
   —mis-application of phonetic patterns
   —poor selection of ambiguous phonetic patterns
   —mis-application of spelling rules and exceptions to spelling rules
3. Errors of production
   —handwriting errors, e.g., "b" for "d"
4. Errors of checking
   —"silly" errors which do not reflect lack of knowledge
   —omissions, transpositions, substitutions

Others have gone a step further in attempts to explain why some people are better spellers than others. Nicholson and Schacter (1979) have identified three kinds of knowledge present in good spellers. First, good spellers tend to have a strong sense of language—a sense of what is right and wrong. Good spellers have a system of internalized rules about spelling. Some of these may have been taught in schools, and some come from experience with language and mental deduction of English spelling patterns. These children, therefore, can predict how to spell words. The third knowledge good spellers possess is visual memory. Because of their keen system of internalized rules about spelling, good spellers can identify those words that violate the rules and simply learn to memorize the unpredictable words.

Walker (1974) studied the visual memory aspect of spelling and concluded it to be the single most significant variable distinguishing good and poor spellers. One hundred forty-six students were given the Mental Imagery Test. This test called subjects to remember geometric images for certain periods of time and bring them to mind upon demand. The same subjects were then given 108 "demon" spelling words to write. Errors in spelling were classified as "P" (due to faulty pronunciation or inappropriate phonic generalization) or "V" (due to unexplained causes). The "V" errors were assumed to be on words that did not follow any rule pattern and, therefore, had to be memorized. The results indicated that females made significantly fewer errors of both Types and that good visualizers of both sexes on the Mental Imagery Test made significantly fewer "V" type errors. Walker proposes that very good spellers may possess a genetic gift for remembering graphic symbols. Once again, it seems that it is helpful to choose your parents carefully, hardly a helpful learning principle.

Other studies indicate that students internalize an efficient system of personal spelling rules. The developmental aspect of spelling was studied in three pieces

of research. By showing how spelling changes over time, Beers and Henderson (1977) confirmed the notion that children build an internalized system of rules that are sophisticated and become effective very early. Early first-grade attempts at spelling rely on a letter-name strategy. Later refined vowel knowledge and orthographic pattern knowledge begins to improve spelling before any formal rules instruction takes place. It would appear that exposure to language improves spelling. Schwartz and Doehring (1977) lend further credence to this theory. They looked at 20 good and 20 poor spellers in each of grades 2 through 5. The children were asked to spell nonsense words which would test their knowledge of orthographic patterns. Clearly it was demonstrated that spelling is developmental in nature; i.e., older spellers performed better even "before the beginning of formal spelling instruction." Further, it was shown that ". . . poor spellers lagged behind the good spellers in pattern acquisition by about two years."

Zutell (1978) also found that children learn to spell by themselves through exposure. His work indicates that children form letter strategies very early in their school career and often over-generalize their internalized rules until repeated exposure and instructional feedback refine their spelling toward conventional usage.

Templeton's (1979) work coroborates Zutell's and shows that children ". . . do not always expect a one-symbol one sound correspondence."

Among the most-blamed causes of poor spelling has been faulty pronunciation or dialectic influences on oral language. Groff (1973, 1978) discusses and rejects much of the previous research in dialect, influences—particularly black inner-city—on spelling. He claims that most of the good research in this area was done with very young subjects. Because it is established that young spellers of all dialects rely on a sound-symbol spelling method until exposure refines their internalized patterns, Groff believes that the influences of dialect are lessened over time. His studies did, in fact, indicate that by the middle and upper elementary grades dialect-related influences in spelling errors were essentially gone and that over time dialect does not interfere with good spelling.

Vocabulary

A dozen or so lists of children's vocabulary are currently available for use by teachers and publishers (e.g., Thorndike, Rinsland, Gates, Dolch, etc.). Most of the lists were constructed many years ago, and because they are outdated have been criticized. Common criticisms suggest that today's child—through travel, television, and tradebooks—is a more sophisticated creature than his or her 1940's counterpart. The older lists do not reflect the many new words that are commonly part of today's school age vocabulary.

Johnson and Majere (1977) revised the Dolch list and published a new basic sight word list of 306 words they feel are most commonly used in today's reading series. Rhode (1977) added a new dimension to her list; by offering many new phrases and technological terms that she believes are part of oral vocabularies of children and should, therefore, be in reading vocabularies. Her modern list includes such terms as "cheeseburger" and "pantsuit."

The construction of vocabularies has been consistent from the early lists to the current ones. Vocabulary items are selected by looking at children's literature: basal reading series, trade books, dictionaries, and classroom materials

of various kinds. Usually with the aid of computers, these words are sorted and edited according to frequency of use. Some lists further delineate the words by the age at which they are typically added to a child's vocabulary.

Word Lists

Historically, word counts and word lists can be traced to the tenth and fifteenth centuries (de Rocher, Miron and Patton, 1973). Those early attempts were predominately biblical concordances. However, by 1721, Nathaniel Bailey had compiled an extensive listing of the English vocabulary which was to serve as the basis of a dictionary. At the turn of the nineteenth century, the Rev. J. Knowles listed 350 most frequently used words from selected literature. Again, however, Knowles' primary source was the Bible and biblical literature.

Early word lists were typically used to specify appropriate spelling lists for school children. Chancellor (1910), for example, compiled a list of some 20,000 words from dictionaries and spellers. He then identified a set of 1,000 words which he deemed most important to teaching spelling to school children. Chancellor did not, however, provide any empirical base for his decisions or any documentation to justify his choices.

An early mode of selecting words was the use of adult writing samples. Ayres (1913) found 2,001 different words used in personal and business letters that he then proposed to serve as a base for spelling instruction. How he selected those letters, however, is unclear.

In another study using correspondence, Andersen (1921) asked his students to bring in personal letters received by parents and friends. He then compiled a corpus of 9,223 words, of which 3,217 had a frequency of one. His entire corpus yielded only 126 different misspelled words. Quite clearly this sampling procedure would lead to a biased sample since the letters produced by the students were most likely unimaginative and certainly not terribly personal. Furthermore, spelling errors would likely be minimized due to personal editing prior to mailing.

In another study, Clarke (1921) analyzed letters to the editor that appeared in Chicago newspapers. The resulting corpus yielded a total of 28,292 running words (Tokens) and 3,360 different words (Types) from 2,000 such letters. A large number of words that appeared in the Clarke list were missing from the Ayres list and vice versa. Both lists reflect biased sources.

The landmark research of Thorndike (1921; Thorndike and Lorge, 1944) dominated the study of reading vocabulary for years and became the keystone upon which later studies were built. Thorndike and Lorge compiled word lists based on children's and adults' reading material. The resulting volumes served as the foundation for vocabulary in reading series and for experimental materials in research on verbal processes for the next fifty years.

The more recent compilation of vocabulary based on children's literature by Carroll and his associates (Carroll, Davies and Richman, 1971) and the analysis of the Brown University corpus of 1,000,000 words (an adult corpus) by Kucera and Francis (1967) moved the study of vocabulary lists forward substantially.

The Carroll, Davies, and Richman (1971) compilation of the "Word Frequency Book" was initiated to identify an appropriate corpus of words for the *American Heritage School Dictionary*. Those researchers identified target reading materials through a survey sent to schools across the nation. Those target materials were then sorted into 22 subject areas, and 200 word samples were taken from

each. The result was a computer-assembled listing of 5,088,721 Tokens (total words) which were sorted into 86,741 Types (different words). Most analyses performed by Carroll, Davies and Richman followed Herdan's (1960) lognormal model. However, like the Thorndike and Lorge word counts, the materials reflect written material that children may encounter rather than that which they produce.

A particularly cogent study from the vantage point of the present research, was conducted by Rinsland (1945). Rinsland compiled an extensive set of children's compositions. In 1936, under funding from the Works Projects Administration, Rinsland had each word from every composition copied onto a file card. These file cards were then sorted and tabulated. The result was the first major analysis of children's written vocabulary. In Rinsland's study, slightly more than 100,000 children's writing samples were analyzed. The result of this effort was the accumulation of 6,012,359 running words and 25,632 different words. Of the 25,632 different words, 11,061, or 43.2 percent, had frequencies of one or two and were not in the published volume. Still, as with other lexicons, the large number of words with a frequency of 1 or 2 accounted for a small portion of the total running words.

In a more recent analysis of children's written vocabulary, Hillerich (1978) generated a list of words based on the compositions of children in one small Illinois school system. Hillerich's list suffers, however, from at least three perspectives. First, the failure to sample from a more heterogeneous population, both geographically and demographically, limits the generality of his findings. While this criticism may be leveled at nearly any study since true random selection is unfeasible, the problem is particularly relevant when samples of geographic convenience are used. Second, Hillerich encouraged teacher corrections and editing of the children's compositions prior to entering the data for analyses. The degree of contamination of results from such interference is immeasurable. Without question, no worthwhile generalizations regarding spelling error patterns would have been possible from such a study. Third, Hillerich provides no analyses of the list of the children's compositions beyond some superficial summary statistics.

VOCABULARY AND READING COMPREHENSION

There is general acceptance that vocabulary facilitates comprehension (Davis, 1944; Ruddel, 1969), that is, the understanding of the sentence or paragraph is enhanced by the knowledge of component words. In other words, vocabulary development and language comprehension are components of the same topic (MaGinitie, 1976). By teaching the skills necessary to determine word meaning, a direct improvement in comprehending related content can be realized (Yap, 1979).

A number of studies that verified the impact of vocabulary instruction on comprehension yielded improved comprehension at the sentence level, (Ahlfors, 1979; Blanchard, 1979; Jenkins, 1978). These studies provided vocabulary instruction of some sort and tested its impact on standardized tests of reading achievement in comprehension. A study by Thompson (1973) produced positive effects in both vocabulary and comprehension after a short period of vocabulary instruction.

A number of studies point to the finding that vocabulary building is in fact concept development. These studies involve some instructional activity in concept development, or some experience related to words that links them to ideas

or concepts (Kaplan & Tuchman, 1980). The result of such efforts shows enhanced concept retention and concurrent vocabulary development. Lieberman (1965) demonstrated this finding at the elementary level, and Dea (1978) and Kessler (1976) showed the increase in concept learnings in secondary history and science classrooms respectively. Humes (1977) demonstrated the effectiveness of using concept-learning techniques to teach sophisticated vocabulary content.

Another strand of vocabulary research relates vocabulary acquisition to classification activities. The readers may classify concepts (Klausmeier, 1974) or may enhance their grasp of a given concept by allowing contexts to build up around a word. O'Rourke (1974) sees vocabulary development as a process of acquiring structures that allow readers to see words as classified components of a synergistic whole. Good readers classify and order concepts (vocabulary). Dillon (1976) documents the continued development and refinement of word classifications across the school grades, and Evanechko and Maguire (1972) add that a substantial change in the organization of word categories occurs as readers mature. It appears that vocabulary development is an integral part of concept acquisition and associated classifications, intellectual operations intimately related to thinking and comprehending.

A position that has been sustained over time holds that vocabulary is a powerful contributing factor to comprehension (Davis, 1944; Thorndike, 1973). This relationship was explored as a causal one by Yap (1979). He analyzed a set of reading data from second- and third-grade children on the hypothesis that vocabulary and comprehension were causally related and that vocabulary was the predominant causal factor. His five procedures for conducting causal analysis yielded considerable convergent validity. The data supported the hypothesis that vocabulary causes comprehension—a significant contribution to this position.

Evaluating Compositions

The study reported in this paper used an evaluation scale to estimate the communicativeness of each composition as a way of relating vocabulary production to the primary purpose for writing—communication.

How to evaluate compositions remains the critical stumbling block in research. All the methods in current use can be divided into two basic Types: holistic and atomistic. Holistic methods treat the composition as a whole entity, whereas atomistic procedures analyze particular aspects of writing.

Holistic Evaluation

Holistic evaluation methods look at compositions as whole messages and rate them quickly and impressionistically. The rater is not interested in counting features of the writing and makes no corrections or tabulations of errors. There are three basic kinds of holistic procedures:

1. Methods that match the student composition with other essays in a graduated series.
2. Methods that assign point values for various general charateristics and derive a total score.
3. Methods that assign an overall letter or number based on a predetermined scale.

Cooper (1977) enumerates six specific holistic evaluation scales, all variations of the three basic Types:

1. Essay Scale: A series of sample essays is offered to which raters match the student composition to the sample most like it. Cooper claims this is probably the least reliable of the holistic methods.
2. Analytic Scale: lists characteristics and a range of scores (e.g., 1 = low, 2 = average, 3 = high) for each quality. The kinds of characteristics typically will include features like organization of composition, theme, vocabulary diversity, etc. When detailed explanation of the characteristics is provided and the raters are trained and practiced Cooper believes this is the most reliable system.
3. Dichotomous Scale: a series of characteristics or questions about the composition to which the rater answers "yes" or "no." The total of "yes" answers is the rating for the composition. Cooper provides no reliability data for this kind of system, but believes it probably is not a very acurate measure.
4. Feature Analysis: focuses on structural and stylistic characteristics of a composition. The rating scale lists qualities related to this one feature. An example of this would be if a researcher wanted to study the style of children's narratives, the scale might include characteristics like emotional quality, imagery and vocabulary variety. This has the same strengths of other analytic scales, but is used only for specific kinds of studies.
5. Primary Trait Scoring: focuses on traits of a composition that are especially related to the writing task at hand, and not for composition in general. For example, the researchers may wish to see how children perform on social correspondence. A task of invitation writing is set up and scored by noting whether the date and place of the party was included correctly, whether a pleasant tone was set in the letter, and so forth. With specific criteria for rating, high reliability can be achieved.
6. General Impression Marking: a scale of usually five to seven increments. The rater assigns an over-all mark from the scale to the paper. This is the simplest of all systems, but requires much rater training and practice before acceptable reliability is achieved.

Another devotee of holistic evaluation, Hillerich (1970), points to his research to support non-syntactic evaluation. He shows why he feels classroom practice and research evaluation of student writing should focus on clarity and interest appeal. He has devised a 6-criteria scale with 1 through 5 points assigned on the qualities: unit of thought, logical order of development, smooth transition, vocabulary variety, sentence variety, and vividness and appropriateness of expression. To this he adds a 9-point mechanical checklist for a general idea about the technical quality of the writing.

Because on the surface holistic evaluation systems appear less "scientific" than atomistic methods, many researchers have addressed the issue of determining how holistic procedures compare to atomistic. Hogan and Mishler (1980) have made comparative studies using both atomistic and holistic methods. They found acceptable reliability for either method: .95 reliability coefficient for holistic methods; .81 to .95 range for atomistic. Their conclusion is that since both methods are about equally effective, one might as well select a holistic method because it is easier and less time-consuming to administer.

Howerton, et al. (1977) studied compositions of 983 subjects in grades 4, 6, 9, and 12. The compositions were rated both quantitatively and qualitatively. The atomistic measures included language productivity, vocabulary diversity, spelling and syntactic maturity. The general quality was judged by the ETS Composition Evaluation Scale. The results demonstrated that quantity and quality are significantly related. Particularly the following atomistic features are related to overall quality: total number of words, total number of sentences, percent of unique words, and number of words per T-unit.

Judine and Griffin (1970) found much the same thing. In a study with 269 seventh, ninth and eleventh graders they discovered that overall quality of compositions was highly correlated to the number of total words, total number of T-units and the total number of clauses.

Ultimately, the quality of holistic evaluation measures rests not with how they compare to atomistic methods but with the quality and training of the raters. They are, of course, the key to the success of holistic evaluation. Experts agree that inter-rater reliability can be in the high .80s and low .90s consistently if raters are trained and have samples representative of the criteria. Cooper (1977) suggests that nearly perfect confidence can be achieved if raters have similar backgrounds and if each student writing sample is rated by more than one rater. Diederich (1974) shows that untrained raters have a .31 reliability. But with practice, reliability rises dramatically. He goes further to recommend a five-point scale: 1 = poorest 5% of papers; 2 = 20% below average; 3 = 50% middle or average papers; 4 = 40% above average papers; and 5 = the top 5% of the papers. Because the most interrater disagreement is found in the middle range, Diederich promises that this scale will render the best results beause it lumps the vague "average" papers into just one category.

Procedures for Collecting and Compiling Children's Compositions

In the fall of 1979, inquiries were mailed to several hundred school systems in the continental United States asking them to participate in a study of children's compositions. In the cover letter, the recipient was informed that the primary purpose of the study was to examine vocabulary patterns of children in grades 1 through 8. The schools were also informed that the compositions would be used to conduct an analysis of other features of children's writing such as organization, theme, sense of punctuation, spelling, capitalization, etc.

This procedure provided a quick sampling of potential school system participants. Each participating school system identified a central contact person who would distribute information and make local arrangements.

A packet sent to each contact person contained five items: a cover letter describing the purpose of the study and the type of participation requested, a set of instructions to teachers, sample instructions for students, example stimuli for student compositions, and a postal reply card. As an inducement to participate, schools were offered a computerized analysis of vocabulary generated by children in that school system.

When a postal card was returned, indicating a school's continuing willingness to participate in the study, a package of materials was mailed. The package included a letter to the teacher, instruction sheets, student data sheets, and mailing labels for return of the compositions. Enough copies of the items were sent to provide one set for each participating classroom.

In January, 1980, a follow-up letter was sent to any school system that had not responded to the original request. If no response resulted from this second request, the school was dropped from the mailing list.

Instructions To Teachers.

Teachers were asked to fill out an information sheet for their class. The information sheet asked for limited demographic data on each of the students and the teacher's subjective rating of the student's reading ability. The teacher was informed that the data were to be used to aid in analyzing the compositions and that all information would be held in confidence, only an identification number would be used and its correspondence to the student's name would be destroyed as soon as the composition had been coded and archived.

Teachers were instructed to ask children to write a composition and to give them 45 minutes to an hour to complete the task. Children were to be encouraged to write about things related to their own lives, especially what they were learning in school, or the people and ideas that they often thought about. On a separate sheet, sample stimuli for compositions were provided along with the suggestion that the teacher might use one or more as needed to stimulate the children's writing.

As children wrote, they were to do so without consulting friends or teachers. Students were to be informed that this was not a test, so they were to spell, punctuate, etc. as they thought those items should appear. Students were encouraged to write clearly and legibly since the compositions were to be read and typed into a computer for analysis.

As soon as the compositions were completed, they were to be packaged securely and returned to the investigators. In all, more than 15,000 papers were received representing widely scattered areas of the United States. Of these, more than 4,000 were included in their entirety for the analysis reported herein. The result was a corpus of nearly one-half million running words (Tokens) which comprise the list described in this study.

Preparation for Analysis.

The student compositions and demographic data were entered directly into permanent files on the Indiana University Wrubel Computing Center system via interactive terminals. Demographic data and a qualitative holistic rating (described elsewhere in this report) were entered first. Eight demographic descriptors included for each child:

1 Identification code.
   a. school system number: 01 to 99.
   b. teacher number, within school: 01 to 99.
   c. student number, within teacher: 01 to 99.
2 Sex of student.
3 Ethnicity.
4 Grade level.

5 Teacher's rating of general reading ability.
6 Bilingual status of the student.
7 Standard Metropolitan Statistical Area (SMSA) rating of the locale of the school system.
8 Communicativeness rating of the composition.

The actual composition was entered directly into computer storage as verbatim text with spelling errors noted and transformed. Capitalization, punctuation and grammar were not altered. Spelling errors were entered in the following fashion: immediately following a spelling error, an asterisk or star (*) was printed. This notation "alerted" the analytic program that the configuration was a misspelled word. The starred misspelling was followed by the corrected spelling enclosed in plus (+) signs. For consensus of spelling, all possible misspellings were compared to the listed spelling in the Macmillan School Dictionary (1977). For example, if a child wrote; Dont forget yor luch mony.
the entered text would be:

Dont* +don't+ forget yor* +your+ luch*
+lunch+ mony* +money+.

Note that the period at the end of the sentence follows the correction of the last word, it did not follow the uncorrected form.

Raters were also informed not to tamper with grammar as long as the word was spelled correctly, that is, if the basic word was a valid word but was used in a manner that might not correspond to standard English, it was not altered. Raters were also given the following guidelines:

Verb forms. Do not make a correction if the basic word is spelled correctly but is wrong in tense or number (e.g., "I see him yesterday." or "The man are not here.") If the student attempts to use a contraction but omits the apostrophe, this *should* be marked as an error in spelling (e. g., "I dont know." would become "I dont* +don't+ know.") If the verb is spelled incorrectly, then make your correction in terms of the verb form the student seemed to be trying to use; this may be difficult to determine in early grades. If the student has written "I sea him yesterday." then use the corrected spelling "see" without altering the tense of the word. If, however, the student attempts something resembling past tense (at least phonetically) then take this into account in your corrections (e.g., "I sow him" or "I seed him" becomes "I saw him" or "The balloon popt" should be coded as "The balloon popped."or "I runned" would be "I ran.")

Noun endings. If the word is spelled correctly but lacks the proper ending (e.g., the singular is used where the plural is intended), do not mark this as an error (e.g., "All the boy are here.") Possessive endings should be marked as errors *if* the student uses an "s" without the apostrophe ("The boys book"). should be coded "The boy's book").) If the "s" is not used, do not add it (e. g., "The boy book" should not be corrected). If the noun is spelled incorrectly or is a non-word, then it should be corrected (e. g., "The bok is on the tabel" or "All the mans are here.") Sometimes a noun or pronoun used is so far off that a basic change has to be made in the form during correction (e. g., "He or she should clean up hes or shes room" cannot be improved with apostrophes and should be altered to "his or her").

Adjectives used where an adverb is required. If the "ly" ending is omitted but the word is spelled correctly, do not mark an error (e.g., "Your work should be done neat").

Contextual errors. In any case which does not clearly fall into one of the categories mentioned above, a decision must be made based on context and meaning and the assumed word the student intended to use. All misspellings of this type should be corrected. Often it is necessary to sound out a word so that a phonetic approximation can be translated into a real word ("clous" = "clothes", "wode" = "would", etc.). Even if the student uses a word which is a legitimate word in its own right, but is wrong in context, it should be marked as a spelling error (e. g., "Timmy want to the circus" or "They came to are house.")

As noted above, raters and recorders of the passages all used the same dictionary. In all error cases, the first spelling was used as the correct spelling. If, however, a student used a secondary spelling, it was rated as acceptable.

In cases where the string of letters was unrecognizable or was perhaps a random string of letters (one such string was 35 letters long) the configuration was coded with a double asterisk (**) immediately following the non-word. The analytic program was alerted to ignore the configuration. Proper nouns were similarly coded with double asterisks to eliminate them from the lexicon. In retrospect, this was not the best decision since it may have been valuable to code non-words and proper nouns differently.

In cases where a compound word was written as two separate words (e. g., "Beat Nik"), the misspelling was enclosed in brackets and the correction was enclosed in the standard plus signs.

The compositions were entered into the computer via text mode over an interactive terminal. The compositions were entered as continuous text in the last 72 columns of a given line of text file. The first 8 columns contained the composition identification code and a line sequence number. The last word in the composition was followed by a double colon (::). This signified the end of the student's composition. If the identification code changed before the double colon was read, an error message was registered by the computer.

Presented below is a sample composition as it was coded:

050103 15120 2063

He kold* +called+ a halcd* +helicopter+ he got in the halcd* +helicopter+ he sed* +said+ will you be mi* +my+ fed* +friend+ Tucdan** and he sed* +said+ bi* +bye+ to his fenz* +friends+ I sed* +said+ gbiy* +good-by+ to* +too+ ::

Rating the Composition.

Prior to its being entered into permanent computer storage, each composition was rated on a scale from 1 to 5 indicating the relative value of communicativeness. The progressive scale, adapted from Smith (1967), attempts to quantify the qualitative differences in children's ability to use language to convey a complete message. As such, it is a holistic rating. The score is not a measure of spelling or grammar although it is likely that those measures are correlated. The skills implied in the rating include (a) a clear conception of words as units of a sentence; (b) an arrangement of words in subject-predicate-object order to form normal sentence patterns or complete utterances; (c)the ability to mark transitions through connectives, such as *and, but, because, then,* etc.; (d) having in mind an overview of a message that should be completed. Given these attributes, each composition was then rated as follows:

1—No communication related to the topic or task beyond one simple utterance (sentence).
2—Communication related to the topic involving several sentences (or equivalent utterances such as run-on clauses) but lacking evident organization.
3—Communication related to the topic involving several cohesive (organized) sentences but lacking completeness (too brief) or are confusing.
4—Communication related to the topic involving a number of cohesive sentences. Communication moves from a beginning, develops the idea and concludes. A complete and organized statement.
5—Same as number 4 plus originality of expression or idea, or a polished writing style that adds clarity and personality to the communication.

Since first-grade children are only beginning to acquire these skills, the scale was altered somewhat to reflect their level of experience:

1—No communication at the sentence level related to the story.
2—Only one idea at the sentence level related to the story.
3—More than one idea communicated at the sentence level, but the story is incomplete or confused or difficult to read.
4—Three or more ideas or events are given in sequence, reaching a satisfactory conclusion, although there may be minor lapses in story or in written expression.
5—Competent (as in 4) plus a variety or sophistication of sentence structure, or originality in bringing about a conclusion.

Raters were admonished to bear in mind that spelling, grammar, and neatness were not to be directly considered in this general measure of communication. These were described as elements that contribute to effective communication, but their perfection is not needed to communicate well.

To merit a rating of 4 or 5 the child must have demonstrated that he knows how to frame a sentence with a capital and a period, even though his composition may have been one long run-on sentence. To get a rating of 4 he should have shown that he has captured the basic idea that written language is expressed in the sentence frame, although he may not have applied the sentence frame to every main subject-predicate-object unit he wrote.

To merit a rating of 5 the child must have used a sentence frame more than once. But again, complete accuracy could not be expected from children whose knowledge of punctuating sentences may have been acquired primarily through an inductive process.

Computerized Analysis.

Once the compositions were prepared, rated, and archived onto permanent 9-track magnetic tape, they were subjected to a sequential set of FORTRAN programs written specifically for this project. The first of the programs transformed the child's composition into a serial alphabetized list of words that appeared in the composition together with coordinate spelling errors where they occurred. The program also generated a data record for each child, which included the demographic data described above, the rating of the composition, the total number of words (Tokens), the number of different words (Types), the total number of spelling errors, and the number of different spelling errors

generated by the child. Those data records form the basis of several of the analyses which follow.

Each of the serial alphabetized children's lists were then merged into a by-school, word-by-grade matrix containing word frequencies in the cells. This by-school vocabulary list was then sent to the participating school. Subsequently by-school lists were then merged into the master list. A complete (300 page) listing is available from the authors. A listing of the 500 most frequent words is attached as Appendix A.

A comparable pattern of merging student lists into by-school lists which were in turn merged into a master list was followed in the generation of a list of spelling errors in the children's compositions. That master list of spelling errors (a 200 page listing) is available from the authors. A listing of the most frequent spelling errors is attached as Appendix B.

The master word list was also merged with a selected set of other lists to provide a basis for analysis of comparability. The lists included the Rinsland (1945) list, the Iowa (Horn, 1926) list and the DiVesta and Walls (1970) children's semantic differential ratings of selected words. Word lists from seven basal reading and spelling series were also merged (Ginn Reading, Scott Foresman Reading, Macmillan Reading, Houghton Mifflin Reading, Holt, Rinehart and Winston Reading, Scott Foresman Spelling, and McGraw-Hill Spelling). This super list showing the relationship among 11 different word lists is available on a selective basis from the authors.

Qualitative Analyses

The master list was matched against other lists to examine the number of differences among the 5,000 most frequently used words. Thus cultural shifts and other qualitative differences could be noted across the time period from Horn (1926) to Rinsland (1945) to the present study. Quantitative differences are also interesting in light of the significance placed on word lists in curriculum development activities.

Analyses of The Pupils' Written Vocabulary List

In all, pupils in this sample produced a corpus of 482,487 running words (Tokens) which were sorted into a list of 10,265 different words (Types). As in previous studies, a few Types accounted for a large portion of language usage. In studies by Cook and O'Shea (1914), Kucera and Francis (1967), Carroll et al., (1971) and Hillerich (1978), for example, the first five or ten Types accounted for upwards to 26 % of the entire corpus of words. Hillerich (1978) found that the first five most frequent words in a study of children's vocabulary accounted for 18.2 % of the entire corpus of written compositions.

TABLE 1

The Ten Most Frequent Types and Their Frequencies (F), Cumulative Frequencies (CF), Proportions of all Tokens (P), and Cumulative Proportion of all Tokens (CP)

Rank	*Type*	*F*	*CF*	*P*	*CP*
1	A	39,599	39,599	0.082	0.082
2	THE	22,599	61,934	0.046	0.128
3	AND	20,496	82,430	0.043	0.171
4	I	12,846	95,276	0.027	0.198
5	TO	11,771	107,047	0.024	0.222
6	WAS	6,673	113,720	0.014	0.236
7	MY	6,363	120,083	0.013	0.249
8	OF	5,855	125,938	0.012	0.261
9	WE	5,514	131,452	0.011	0.271
10	HE	5,435	136,887	0.011	0.284

In the same fashion, a few Types dominated children's writing in this study. Table 1 presents the 10 most frequent Types in the resulting list in descending order of their frequencies (f). Also presented are cumulative frequencies (cf), proportion of all Tokens (p) and cumulative proportion of all Tokens (cp). It may be seen in that table that the first five Types accounted for 22.2 % of the entire corpus. The second five Types only accounted for an additional 6.2 % of the Tokens. (The first 500 most frequent words are presented in descending order of frequency as Appendix A.) By way of contrast, at the other extreme, 3,550 Types had a frequency of 1. Hence, in the low frequency end of the distribution, 34.6 % of all Types accounted for less than 1.0 % (0.7 %) of the Tokens. When Types with a frequency of 2 are included, another 1,531 Types are added. In so doing, roughly 50 % of the Types provide a combined contribution of only 1.4 % of the entire corpus of Tokens.

Distribution Density

Kucera and Francis (1967) and Herdan (1960) have noted that an additional index of the density of cases within a few Types may be gained through the use of Yule's *K*. *K* provides a value linked to distribution values and increases as the distribution becomes increasingly skewed (Yule, 1944). In those Type-Token distributions that display a heavy dependence on a few Types, values of *K* are elevated. In Kucera and Francis's analysis of the distribution density of the Brown University corpus of adult reading material, for example, they reported a value of $K = 98.7077$ which they took to indicate a very heavy density. That is, most adult reading material was built around a minority of Types. Previous studies which report a higher concentration of common words in children's productive vocabulary, would indicate that higher values of *K* would indicate result from a corpus of children's writing.

Table 2 presents a set of distribution statistics for the lists by grades 1 to 8 and for the entire corpus. The values of *K* range from a low of 142.728 at the

second grade to a high of 160.568 at the fourth grade. There was no regular shift in values of *K* across grade levels. Clearly, however, the density distribution was more extreme in the children's compositions than in adult reading matter. Whether an analysis of adult written performance would yield greater or lesser values of *K* is unclear at this point, however, one would suspect a reduction of those values with improved lexical flexibility.

TABLE 2

Distribution Statistics of the List by Grade and by the Overall List

Gr	*Mean* T/T Ratio	*Standard Deviation*	*Skew*	*Kurtosis*	*Yule's K*	*Types*	*Tokens*
1	10.323	36.756	10.993	152.087	160.121	848	8754
2	13.508	67.992	16.644	340.871	142.728	1840	24855
3	16.229	105.816	22.349	612.380	159.982	2716	44078
4	18.495	143.613	27.331	944.123	160.568	3814	70540
5	17.633	142.867	28.593	1044.893	151.479	4396	77513
6	18.360	156.980	31.544	1285.034	147.657	5015	92076
7	17.031	151.829	37.008	1788.169	155.424	5174	88120
8	15.359	135.054	29.947	1092.967	157.010	4984	76551
	47.003	574.298	45.005	2618.645	146.387	10265	482487

Analyses of Children's Compositions

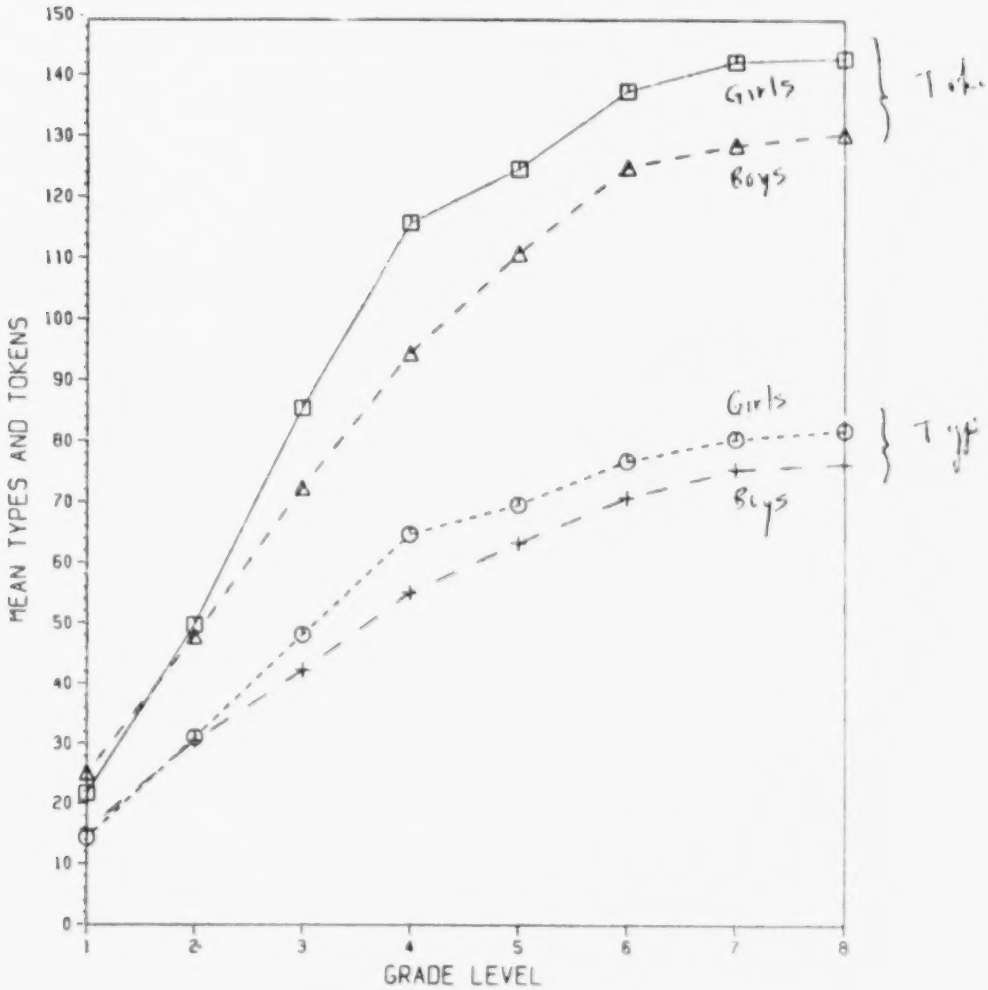
In a study of predictive validity between multiple choice tests and holistic ratings of children's compositions, Hogan and Mishler (1980) found correlations of 0.65 for third graders and 0.68 for eighth graders. Lindell (1980), reporting on a Swedish study of children's writing behavior, noted that the single best predictor of a holistic score was the number of Types (different words) in the composition. Further, other mechanistic attributes of the composition such as punctuation, the number of lines, the number of Tokens (total words), different adverbs were also significant predictors of holistic values assigned to Swedish children's compositions.

Token Production

Students in this sample wrote compositions containing an average of 102.03 Tokens per composition with a standard deviation of 62.70 Tokens. On the average, girls produced more Tokens per composition (M = 109.10) than did boys (M = 94.72) and the advantage was stable across grades. Figure 1 presents mean token and type productions by composition for boys and girls across the eight grade levels. Only at grades 1 and 2 was some ambiguity regarding sex differences in token production seen. However, no statistically significant interaction of sex and grade was yielded for the data. As would be expected, children produced more Tokens per composition as they progress through the first eight grades of schooling ($p < .001$).

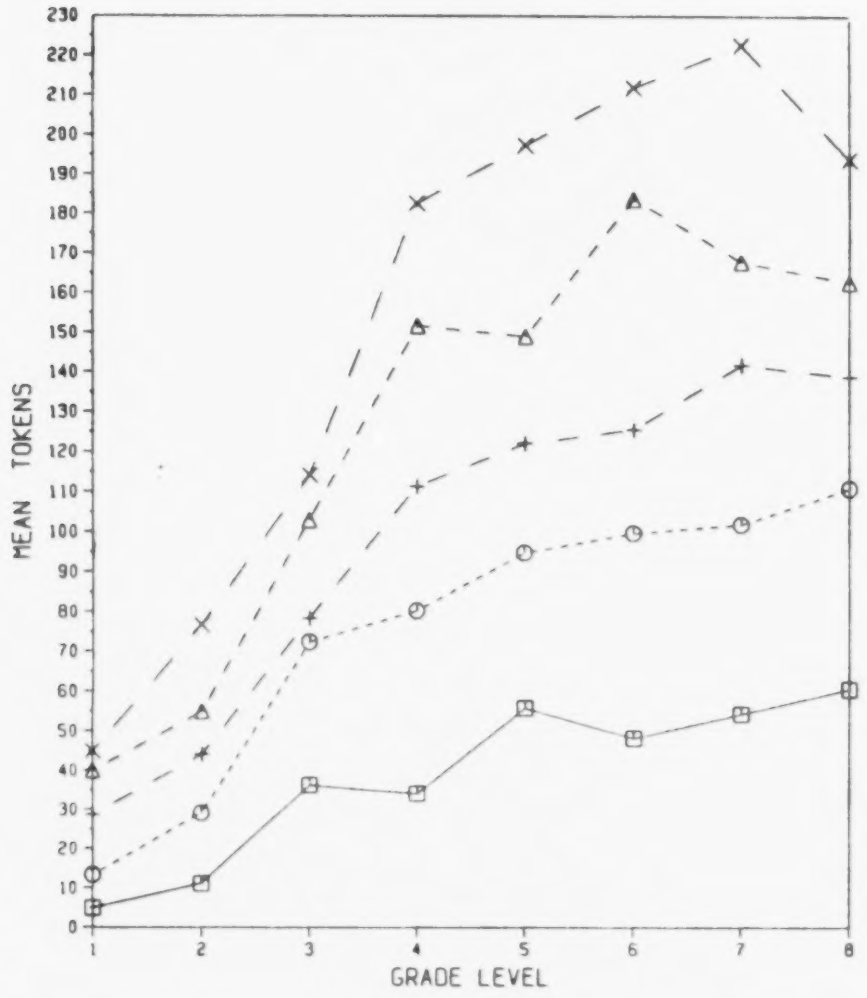
FIGURE 1

TYPE AND TOKEN PRODUCTION BY SEX AND GRADE



Further, increased token production at each grade was higher among those compositions rated higher in communicativeness ($p < .001$). There was, however, a statistically significant interaction of the effects of grade and holistic rating on token production. The grade by rating contribution to token production may be seen in Figure 2. Note in that figure that while higher holistic ratings were linked to higher token production at all grade levels, the advantage became more marked as the grade level increased.

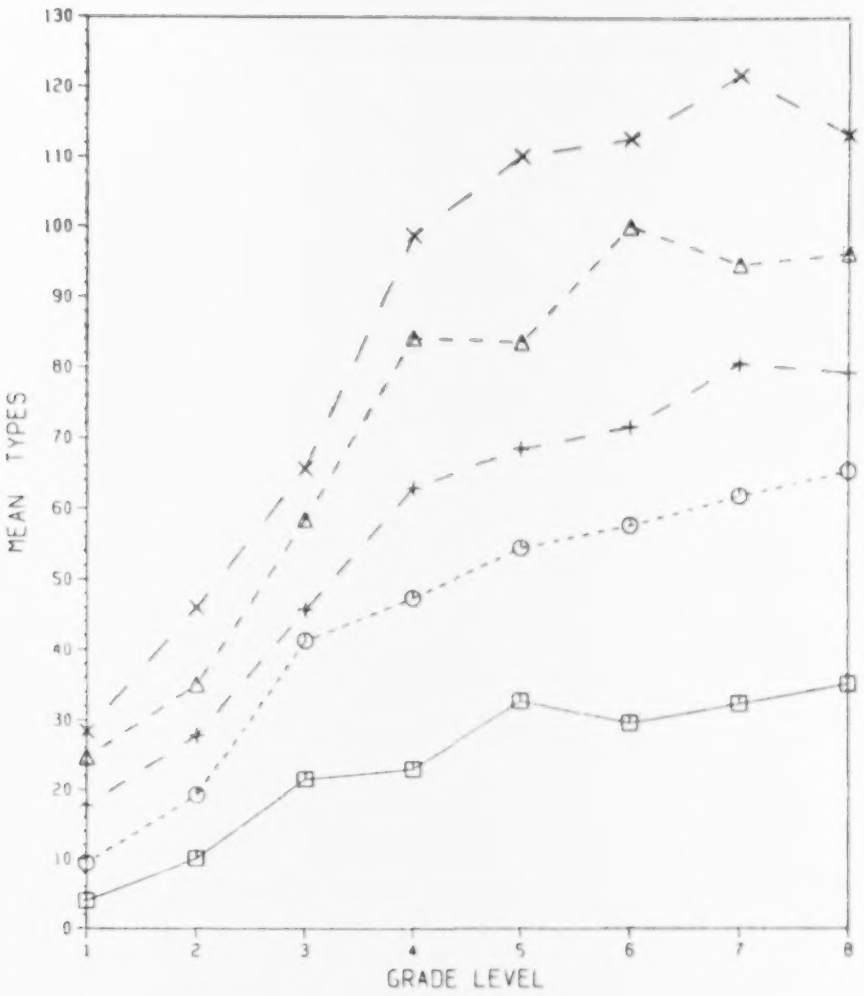
FIGURE 2
TOKEN PRODUCTION BY RATING AND GRADE



Type Production

Students in the study produced an average of 59.25 Types per composition with a standard deviation of 31.87 Types. As with token production, girls were more likely to produce more Types per composition (M = 62.34) than boys (M = 56.04), and type production increased with grade level ($p < .001$). Type production for boys and girls over the eight grade levels may be viewed in Figure 1. Further, as with token production, children whose compositions were rated as higher, more communicative, had more Types per composition ($p < .001$). However, unlike token production, there was not a holistic rating by grade-level interaction. The nature of the relationship between holistic rating, grade level and type production may be seen in Figure 3.

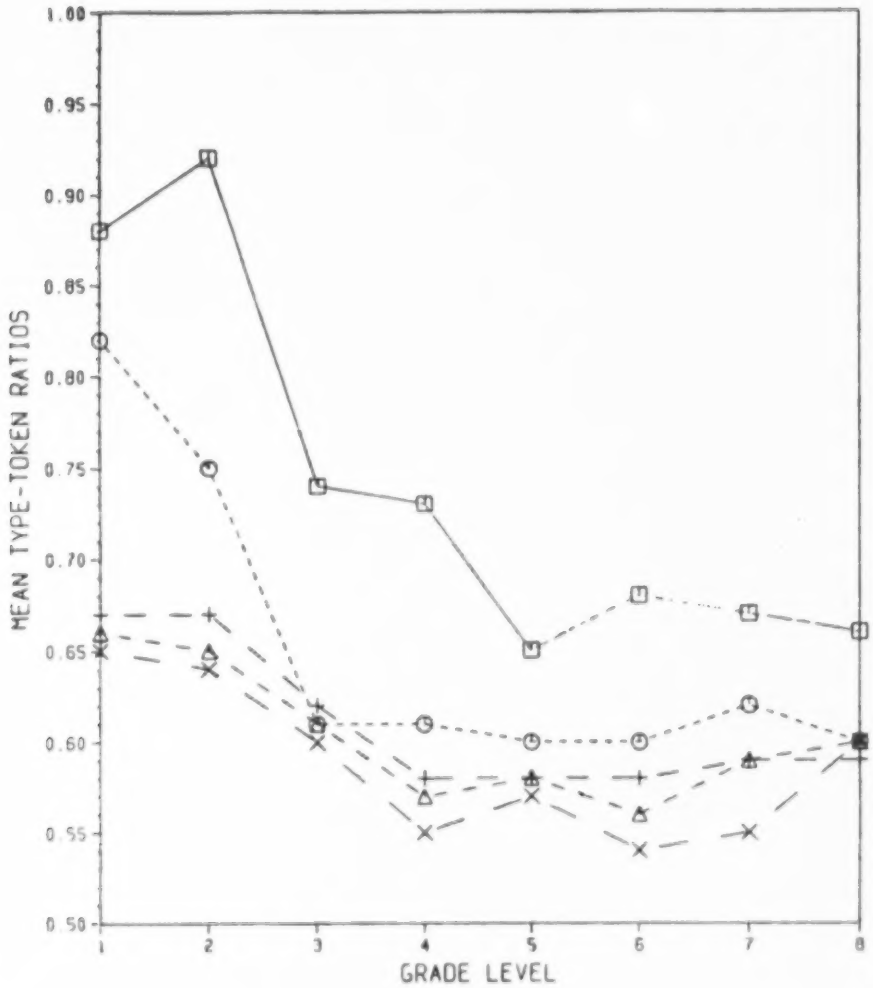
FIGURE 3
TYPE AND PRODUCTION BY RATING AND GRADE



Type-Token Relationships

In early studies of the relationship of type to token production in children's language, a regular decline in the magnitude of the ratio of Types to Tokens was observed as children matured (Chotlos, 1944; Johnson, 1944). While the type-token ratio reflects a degree of diversity of language, the volume of a child's verbal output is directly related to the reliability of the ratio, i.e., as volume of output increases, so does the reliability of type-token relationships (Chotlos, 1944). Since the volume of output increased across grade levels, there is a reasonable expectation that reliability suffered among ratios for younger children. Further, as the correlation between the number of Types and Tokens increases the usefulness of the ratio in correlational studies becomes suspect.

FIGURE 4
TYPE-TOKEN RATIOS BY RATING AND GRADE



Carroll (1968) cautions that a decrease in type-token ratios across age groups may be artifactual because of differing base rates of Tokens. He suggests a modification of the simple type-token relationship by dividing the number of Types by the square root of two times the number of Tokens. While the latter transformation may be mathematically more satisfactory, it fails to yield a psychologically interpretable result. Hence, the cautions defined above nothwithstanding, the following results were observed between type and token production.

There was a strong linear relationship between the number of Types and Tokens produced by children in the study. In assessing the magnitude of the relationship of the unadjusted values of the two, the squared correlation was 0.922. There was a slight nonlinear effect. Following the lead of Chotlos (1944) and Carroll (1968; Carroll, et al., 1971) the type and token values were trans-

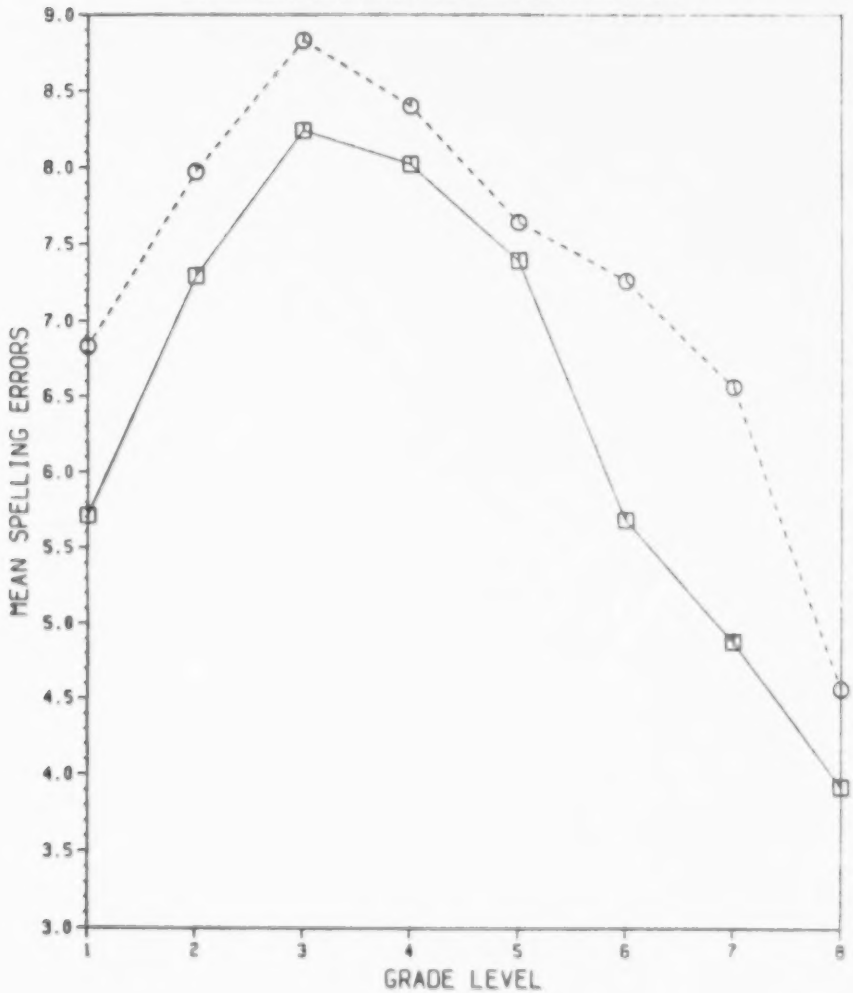
formed into logarithmic equivalents. The squared correlation between the logarithmically transformed type and token productions was 0.933. Clearly there was little room for growth in the zero-order relationship but the increase resulting from the transformation was significant. The regression line resulting from that analysis was

$$\text{Log (Types) } ' = 0.8489 * \text{Log (Tokens)} + 0.1597$$

This function is similar to that reported by Chotlos (1944), although the magnitude of the function in this study was substantially higher. When viewing the same relationship within subgroups defined by holistic ratings of the compositions, the log-linear slopes were roughly parallel and the intercepts increased regularly with the holistic rating.

Finally, Figure 4 presents the mean type-token ratios for children in the study by grade and by holistic rating. As with previous analyses (see especially,

FIGURE 5
SPELLING ERRORS BY SEX AND GRADE



Chotlos, 1944), the magnitude of the ratio diminishes across grades and, to a lesser extent, across holistic ratings.

Spelling Error Production

Students in the study produced a corpus of 10,473 different spelling errors. The average number of spelling errors in a given composition was 6.89 with a standard deviation of 6.91 spelling errors. The distribution, as may be inferred from the above listed data, was positively skewed. The number of spelling errors per composition ranged from 0 to 110. Girls produced more (M = 7.37) than boys (M = 6.42) and number of spelling errors increased initially then diminished with grade level ($P < .001$). This relationship may be seen in Figure 5.

Curiously, children whose compositions were rated higher in communication value produced more spelling errors than those whose compositions were rated as lower in communicativeness ($P < .001$). This, and the apparent disadvantage seen among girls may result from the fact that higher-rated compositions and compositions written by girls had more Tokens. The larger compositions provided greater opportunity for spelling errors. A more appropriate measure would thus be a ratio of the number of spelling errors to the total number of Tokens. When this ratio was created the results were more in accord with expectations. Higher-rated compositions had lower spelling errors to Token ratios than their low-rated counterparts. An exception was seen among first-graders. However, the instability of first grade is partially artifactual since a large proportion of first-grade compositions that were rated as "1" had one or two words that may have been spelled correctly but communicated little. The ratio in those cases would be 0.00. The distribution of the spelling error proportion for compositions of differing ratings across grades is seen in Figure 6.

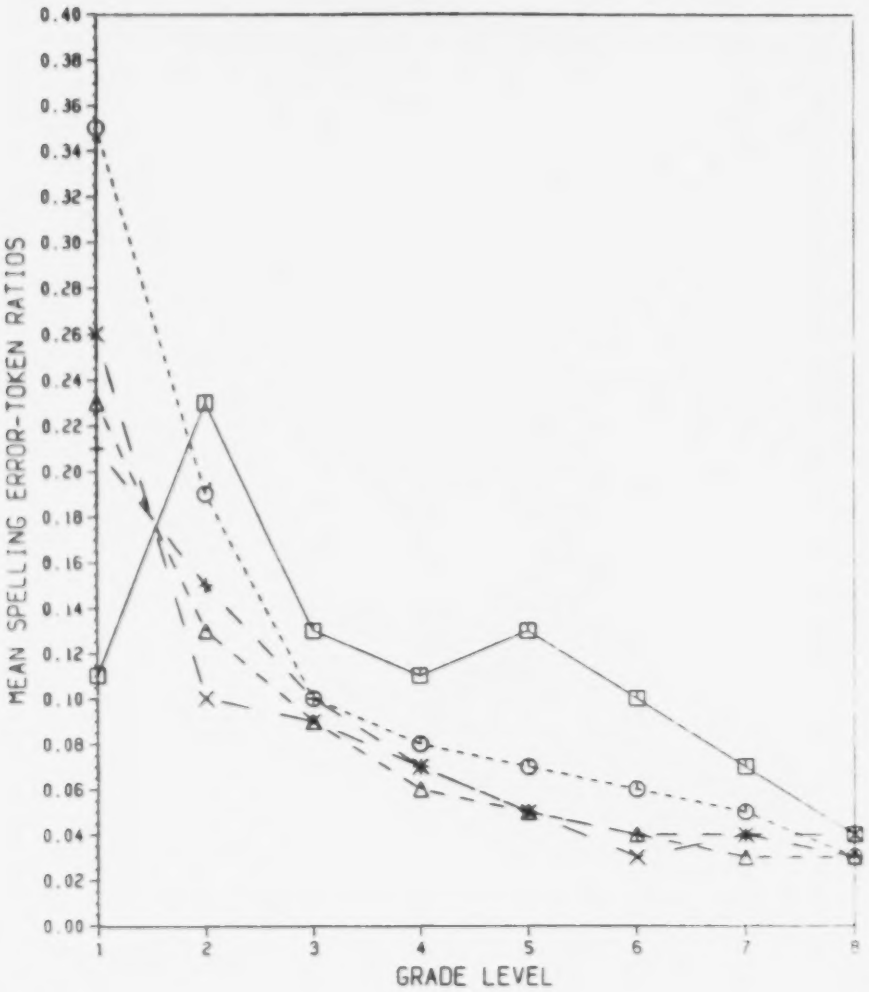
Wing and Baddeley (1980) recently noted the need for a corpus of spelling errors and analyses of such a corpus. They present analyses of spelling errors generated from 40 college students' essays. On the average, college students generated about 1.5 spelling errors per 100 words. In comparison, the rate for eighth-grade children in this study was about 3.5 per hundred words and about 25.0 per hundred words for first-grade children. There was an apparent non-linear decline in error rate from grades 1 to 8. This non-linear function may be viewed from its logarithmic function. The resulting log-linear relationship yields an $R^2 = .209$; the regression line for that relationship is

$$\log (\text{Error Ratio}) = -.960 \log (\text{Grade}) - .574$$

If the regression line is projected out to the sophomore year (Grade = 14), the projected error rate is 0.5 per 100 with a one standard deviation confidence interval of 0 and 1.8, which includes the Wing and Baddeley value.

A limited listing of highly frequent spelling errors is included as Appendix B. The listing is organized by a given word, its common misspelling, and the rate of occurence at each grade level. One should be careful, however, not to assume that these are representative of the entire range of errors. The entire list of 10,473 different spelling errors is available from the authors.

FIGURE 6
SPELLING ERROR-TOKEN RATIOS BY RATING AND GRADE



Vocabulary Changes Over Time

The volume of word production is, of course, a function of several factors, especially a pupil's available vocabulary and practice in writing. The powerful effect of vocabulary on reading comprehension (more than 50% of the variance) has been demonstrated by Davis (1944), Thorndike (1973-74), and Yap (1979). This current composition study shows a similar effect on pupil's writing. The difference in volume and in the number of unique words produced can be illustrated with samples taken from first grade pupils.

Two pupils who scored in the bottom quarter in both reading and written communication wrote the following:

Pupil 1: his dad was not happy.

Pupil 2: I went to grandmas house.

Two first grade pupils who scored in the top quarter in both reading and written communication wrote the following:

Pupil 3: I went roller-skating with Jean and Dawn.
It was fun. There was music. We did the bugie wugie too. My mom can't ice skate so you can imagine what she looks like on wheels.

Pupil 4: Meredith is my best friend. She hits me alot but thats OK we still get along. Sometimes we play baseball socker and football. Sometimes we swim in her blue sement swimming pool.

There are, of course, other factors operating in these examples besides volume of output and the number of unique words. Those factors, however, are not the focus of this paper.

Several comparisons can be made between the vocabulary list generated by the current sample and lists from earlier studies. Most applicable is the Rinsland study (1945) which also generated a vocabulary from pupils' writing, grades one to eight. Some of the differences can be attributed to the cultural shifts that have occurred over the past four years. Sample changes are illustrated below.

CULTURAL

There is a change in the vocabulary that children use, e.g., cultural changes: an ill-mannered person isn't *crude* (1945), *he/she/it* (you guessed it, *gender* is no longer used) is *gross* (1981).

And it's not because of people's *ignorance* (1945); it's because of their *stupidity* (1981).

Some people were a *nuisance;* now they're a *pest.*

Children wouldn't be caught dead in *trousers;* you *gotta* wear *jeans.*

The sermons at churches must be losing their fire: *ministers,* not *preachers,* are giving them today.

It looks as if the following words might be here to stay for a while:

all-star	double-deckers	madder
back-up	four-seater	minibike
beeper	fro	neater
between-meal	glob	neatest
bionic	goof	plop
blast-off	gooey	pro
blob	gory	runt
brand-new	guts	sissy
built-in	guzzlers	slop
burp	he/she (as a unit)	snuck
chalk-talk	hi	ten-speed
chugging	him/her/it (variant	ugh
clobber	of he/she/it)	weirdest
conk	hogging	whoopee
cooped	jerks	wow
disco	licks	

Still think times haven't changed? Elementary school children are using the following words in their writing today: abortion, acid, addicted, assassinate,

brutality, bugging, busing, corruption, dope, downs, drugs, euthanasia, fuel-efficient, ghetto, hassle, heroin, holocaust, junkie, live-in, long-hairs, magnum, marijuana, molested, mugged, no smoking, nuclear, overpopulated, overdose, pollution, panic-stricken, power-mad, prefabricated, push-button, radiation, rape, rerun, riots, robot-like, sex, sexy, sixpack, slums, smog, terrorized, tranquilizers, warheads, and zillionaire. None of these words appeared in children's writing in the forties.

AUDIENCE

Two other comparisons reflect the qualitative shifts represented by this vocabulary study. Between the first 5,000 most frequently used words in the current sample and the first 5,000 in the Rinsland sample there are 1,758 different words. A similar change can be noted when the current sample is compared to the Horn list (1928), a vocabulary list often used by publishers of childrens books, a list generated from writing for adults. A difference of 1,915 words occurs between the first 5,000 on the Horn list and the first 5,000 most frequently used words in the current sample. Appendix A contains the 500 most frequently used words in the current study. The complete word list, a 350-page document is available from the authors.

Conclusions

Beyond the specific analyses and results in this monograph, the data from this study serve as an important resource for future research. First, the lexicon that results from the children's compositions provides a current list on which future evaluative studies of children's writing may be based. Beyond offering a broad template against which to assess the range and variety of children's writing at the elementary levels, it sets the stage for the establishment of criteria for assessing individual productivity. Second, the merged "super list" provides a capability of comparing current writing behavior to other templates. Third, the generation of the compendium of spelling errors for such a broad sample of children's writing behavior may be unique. Insofar as our literature search has yet revealed, no other such compendium exists over the past 50 years. It is clearly an important resource for future spelling research. Fourth, the availability of data on individual compositions provides a base on which other normative and descriptive research may be based. Fifth, the availability of the actual compositions permits future quantitative and qualitative analyses beyond those described in this monograph. For example, future researchers might study organizational patterns, ethnic and regional characteristics, and so on. Sixth, the availability of both quantitative and qualitative data on children's compositions provides a base upon which the relationship of mechanistic and holistic evaluations may be determined.

The analyses reported in this monograph tend to confirm and to extend results of previous research in children's writing. Qualitative and quantitative trends were observed related to maturity and ability. Further, a general picture of children's writing behavior emerged.

In some sense, the picture of current writing performance is less than encouraging. In both quantitative and qualitative analyses, the compositions were disappointing. Quantitatively, most written output was made up of a very small subset of Types. Qualitatively, the compositions lacked spontaneity or originality. By and large, they were boring and banal in the judgment of the raters.

Nonetheless, several useful characteristics of children's writing may be gleaned from these analyses. First, simple quantity of output and quantity of Types output increased regularly over grade levels, with girls holding an advantage in quantity of production. Further, as the holistic qualitative rating of the composition increased, children were more likely to produce more Types and more Tokens and fewer spelling errors. School activity over time pays off in those measures. The more communicative compositions were notable by their quantity and breadth of output and their lower rate of mechanistic errors. A word of caution about the latter relationship is in order. It is not totally certain that type and token production and spelling errors do not contaminate raters' behavior. Greater quantity of output and low spelling errors may create a halo effect. Certainly, previous writing research has indicated this to be so. Nonetheless, these analyses and others have indicated that type and token production are significant indicators of quality of output. If quality of composition is seen to reflect general ability, then the results of these studies parallel those of Chotlos (1944) who found comparable differences in type and token production related to standard tests of intellectual ability.

Whether diversity of output improves with grade level is less clear. Overall lexical diversity of language was noted across grades. In viewing type-token ratios, some improvement was found across grade levels and across qualitative ratings of the compositions. Answers to questions regarding diversity may thus be tied to the manner in which the data are analyzed.

These data certainly indicate some developmental considerations for the elementary school curriculum, grades 1 through 8, ages 6 through 14. Volume and diversity of vocabulary production, major correlates to success in written communication and in reading comprehension, need to be promoted among students, especially those who are less successful in the early years. Spelling errors, related as they are to success in written communication, may be dealt with as specific words or as patterns of words which are needed in a given grade. These data indicate the ages at which students want to use certain words—with their misspellings—and thus enable curriculum developers a more secure sense of the likelihood of success in grade by grade spelling activities.

Finally, this study reminds educators that measures of composition effectiveness may be both qualitative (holistic) and quantitative (mechanistic) with very high correlations between those types of measurement. The data from this study provide educators with several quantitative measures, in addition to a qualitative measure, as bases for examining compositions developmentally from ages six to fourteen.

References

Ahlfors, G. *Learning word meanings: a comparison of three instructional Procedures.* Doctoral dissertation, University of Minnesota, 1979.

Allen, R. *English grammars and English grammar.* New York: Scribner's, 1972.

Anderson, W. N. *Determination of spelling vocabulary based on written correspondence.* Iowa City: University of Iowa Studies in Education, 1921, *2*, No. 1.

Ayres, L. P. *The spelling vocabularies of personal and business letters.* New York: Russell Sage Foundation, 1913.

Beers, J. W. and Henderson, E.H., A Study of developing orthrographic concepts among first graders, *Research in the Teaching of English*, 1977, *2*, 133-148.

Blanchard, J. S. *Word pronunciation and comprehension.* (Doctoral dissertation, University of Georgia, 1979.) *Dissertation Abstracts International*, 1979, *40*, 1979A.

Blount, N. S. Research on teaching literature, language and composition. In R.M.W. Travers (Ed.), *Second Handbook of Research on Teaching*. Chicago: Rand McNally, 1973.

Blount, N.S., Frederick, W.C. and Johnson, S.L. Measures of writing maturity from two 500-word writing samples. Report from the Individually Guided Instruction in English Composition and Literature Project. Madison, Wisc.: University of Wisconsin, 1969. (ERIC No.: ED 036 522.)

Braddock, R.R., Lloyd-Jones, R. and Schoer, L. *Research in written composition,* Champaign, Ill.: National Council of Teachers of English, 1963.

Buros, O.K., (Ed.), *English Tests and Reviews*, Highland Park, N.J., The Gryphon Press, 1975.

Calkins, L.M. Learning to throw away, *Language Arts*, 1979, *56*, 747-752.

Carroll, J.B. On sampling from a log-normal model of word-frequency. In H. Kucera and W.N. Francis, *Computational analysis of present-day American English*. R.I.: Brown University Press, 1967.

Carroll, J.B. *Language and thought.* Englewood Cliffs, N.J.: Prentice Hall, 1968.

Carroll, J.B., Davies, P., and Richman, B. *Word frequency book.* New York: American Heritage, 1971.

Chancellor, W. E. Spelling: 1000 words. *Journal of Education*, 1910, *71*, 488-489, 517-518, 545-546, 573, 578.

Chotlos, J.W. A statistical and comparative analysis of individual written samples. IV: Studies in language behavior. *Psychological Monographs*, 1944, *56*, Whole No. 255, 77-111.

Cook, W.A. and O'Shea, M.V. *The child and his spelling.* Indianapolis, Bobbs-Merrill, 1914.

Cooper, C. R. *Holistic evaluation of writing,* in C.R. Cooper and L. Odell, (eds.) *Evaluating writing: describing, measuring, judging,* Champaign, Ill.: National Council of Teachers of English, 1977.

Cronnell, B. *The scoring of writing samples: A study.* Los Alamitos, Calif: Southwest Regional Laboratory for Educational Research and Development, 1980.

Davis, F. Fundamental factors in comprehension in reading, *Psychometrika*, (September, 1972) 185-97.

Davis, F. Psychometric research on comprehension in reading, *Reading Research Quarterly*, 1972, *7*, 628-678.

Dea, W.A. *The relationship between vocabulary recognition and higher chapter test scores in United States History classes.* Masters thesis, Whittier College, 1978. (ERIC No: ED 159 595)

deRocher, J. E., Miron, M. S., Patton, S. M. and Pratt, C.C. *The counting of words: A review and theory word counts with annotated bibliography.* Syracuse, N.Y.: Syracuse University Research Corporation, 1973. (ERIC No: ED 098 814)

Diederich, P. B. *Measuring growth in English,* Champaign, Ill.: National Council of Teachers of English, 1974.

Dillon, D. Semantic devlopment of selected lexical items as studied through the process of equivalence formation. Paper presented at AERA Convention, San Francisco, 1976.

DiVesta, F.J., and Walls, R.T. A factor analysis of the semantic attributes of 487 words and some relationships to the conceptual behavior of fifth-grade children. *Journal of Educational Psychology,* 1970, *61* (2), Whole No. 6.

Dupuis, M. M. Transformational analysis of compositions, 1972, ED 091-747.

Edwards, T. O. An objective evaluation of expository essays, 1976. (ERIC No.: ED 111 007).

Evanechko, P.O., and Maguire, T.O. The dimensions of children's meaning space. *American Educational Reserch Journal,* 1972, *9,* 507-523.

Fagan, W. T., C. R. Cooper, and J. M. Jensen *Measures for research and evaluation in the language arts,* Champaign, Ill.: National Council of Teachers of English, 1975.

Finn, P. J. Computer-aided description of mature word choices in writing. In C.R. Cooper and L. Odell, (eds.)*Evaluating Writing: Describing, measuring and judging,* Champaign, Ill.: National Council of Teachers of English, 1977.

Follman, J. C., and J. A. Anderson An investigation of the reliability of five procedures for grading English themes. Champaign, Ill.: National Council of Teachers of English. (ERIC No.: ED 016 679.)

Frederick, V. Writing assessment research report: A national survey. Madison, Wisc.: Wisconsin State Department of Public Instruction, 1979.

Freedman, S. W. Models of The evaluation process, paper presented to the annual Conference on College Composition and Communication, Kansas City, Mo., April, 1977. (ERIC No.: ED 146 610.)

Friedman, M., and Fowler E. Assessing elementary students' writing skills. Publication No. 78.74. Austin Independent School District, Texas Office of Research and Evaluation, April, 1979. (ERIC No.: ED 172 201.)

Froese, V. Understanding writing, *Language Arts,* 1978, *55,* 811-815.

Glazer, J. Glazer narrative composition scale, (Doctoral dissertation, The Ohio State University, 1971.) *Dissertation Abstracts International,* 1972, *32,* 3556A. (ERIC No: ED 091 763.)

Graves, D.H. *Balance the basics: Let them write.* New York: Ford Foundation, 1978.

Graves, D.H. A new look at research on writing. In Shirley Haley-Jones (ed.), *Perspectives on writing research in grades 1-8.* Champaign, Ill: National Council of Teachers of English, 1981.

Greenhalgh, C. and Townsend, D. Evaluating students' writing holistically—An alternative approach. *Language Arts,* 1981, *58,* 811-822.

Groff, P., Children's speech errors and their spelling, *Elementary School Journal.* 1973, *74,* 88-96.

Groff, P. Children's spelling of features of black English,: *Research in the Teaching of English,* 1978, *12* (1), 21-28.

Halpern, S., Spreitzer, E. and Givens, S. Who can evaluate writing? *College Composition and Communication.* 1978, *29,* 396-397.

Herdan, G. *Type-token mathematics.* The Hague, Netherlands: Mouton, 1960.

Hillerich, R.L. *A written vocbulary of elementary children.* Springfield, Illinois: C. C. Thomas, 1978.

Hillerich, R. L. Evaluation of written language, Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, March, 1970. (ERIC No.: ED 041 944.)

Hogan, T. P., and Mishler, C.J. *The Elementary School Journal.* 1979, *79,* 142-146.

Hogan, T.P., and Mishler, C. Relationship between essay tests and objective tests of language skills for elementary school students. *Journal of Educational Measurement,* 1980, *17,* 219-227.

Horn, E. *A basic writing vocabulary* (University of Iowa Monographs in Education. First Series No. 4) Iowa City: University of Iowa, 1926.

Howerton, M.L.P., Jacobson, M.D., and Selden, R.W. The relationship between quantitative and qualitative measures of writing skill. Paper presented to Annual Meeting of American Educational Research Association, April, 1977. (ERIC No.: ED 137 416.)

Hudson, S.A. and Veal, L.R. An empirical investigation of direct and indirect measures of writing. Report of the 1980-81 Georgia Competency Based Education Writing Assessment Project. Atlanta, Ga.: Georgia State Department of Education, 1981.

Humes, A.. *The use of concept-learning techniques in vocabulary development instruction.* Los Alamitos, Calif. Southwest Regional Laboratory for Educational Research and Development. 1977.

Hunt, K. W. Early blooming and late blooming syntactic structures. In C.R. Cooper and L. Odell (Eds.,) *Evaluating Writing: Describing, Measuring and Judging.* Champaign, Ill.: National Council of Teachers of English, 1977.

Jenkins, J. R. *Vocabulary and reading comprehension: Instructional effects.* Technical Report No. 100. Urbana: Center for the Study of Reading, 1978.

Johnson, D., and Majere, E. Johnson's basic vocabulary words for grades 1 and 2. *Elementary School Journal,* 1977, *77,* 74-82.

Johnson, W.C.A. Program of research: Sudies in language behavior. *Psychological Monographs,* 1944, *56, Whole No. 255, 1-18.*

Judine, Sister J. M. and Griffin, W.J., Relationships between overall quality and seven language features in composition written in grades seven, nine and eleven. Nashville, Tn.: George Peabody College for Teachers, 1970. (ERIC No.: ED 046 932.)

Kaplan, E. M., and Tuchman, A. Vocabulary strategies belong in the hands of learners. *Journal of Reading.* 1980, *24,* 32, 34.

Kessler, C. The Relationship between Science education and language development. Paper presented at the annual convention of the American Education Research Association Convention. San Francisco, Calif. 1976.

Klausmeier, H.J. and Ghetala, E.S., and Frayer, D.A. *Conceptual learning and development: A cognitive view.* New York: Academic Press. 1974.

Kucera, H. and Francis, W.N. *Computational analysis of present-day American English.* Providence, Rhode Island: Brown University Press, 1967.

Lieberman, J. *The effect of direct instruction in vocabulary concepts on reading achivement.* New York University: Unpublished doctoral dissertation, 1965.

Lindell, E. Six reports on free writing: A summary of the FRIS project. *Didakometry* No. 81. Malmo, Sweden: School of Education, 1980.

Lloyd-Jones, R. Primary trait scoring, in C. R. Cooper and L. Odell, eds., *Evaluating writing: describing, measuring, judging.* National Council of Teachers of English, 1977.

Loban, W. *Language development: Kindergarten through grade twelve.* (Research Report 18.) Urbana, Ill.: National Council of Teachers of English, 1976.

MaGinitie, W. H. Research suggestions from the literature search, *Reading Research Quarterly,* 1976, 2, 7-35.

McCleary, W. J. A note on reliability and validity problems in composition research, *Research in the Teaching of English,* 1979, 13, 274-277.

Macmillan School Dictionary, W.D. Halsey,(ed), New York, Macmillan Publishing Co., Inc. 1977.

Mullis, I.V.S. *The primary trait system for scoring writing tasks.* Denver, Colo: Educational Commission of the States, 1976.

Mullis, I.V.S., and Mellon, J.C. *Guidelines for describing three aspects of writing: Syntax, cohesion and mechanics.* Denver, Colo: Educational Commission of the States, 1980.

National Assessment of Educational Progress. *Writing achievement, 1969-79: Results from the third national writing assessment, Volume III—9-year-olds.* Denver, Colo: Educational Commission of the States, 1979. (a)

National Assessment of Educational Progress. *Writing achievement, 1969-79: Results from the third national writing assessment, Volume II—13-year-olds.* Denver, Colo: Educational Commission of the States, 1979. (b)

National Assessment of Education Progress. *Writing mechanics 1966-1975: A capsule description of changes in writing mechanics.* Report No. 05.W-01. Denver, Co.: National Assessment of Educational Progress, 1975.

Newkirk, T., The Mass Testing of writing: How well is it being done? University of New Hampshire, 1977. (ERIC No.: ED 158 310.)

Nicholson, T., and Schacter, S. Spelling skill and teaching—Putting them back together again. *Language Arts,* 1979, 56, 804-809.

Norris, E. L., et. al. National Assessment of Educational Progress, National Center for Educational Research and Development, 1970, ED 051-245.

O'Donnell, R. C., Griffin, W. J., and Norris, R. C. *Syntax of Kindergarten and Elementary School Children: A Transformational Analysis.* (Research Report.) Champaign, Ill.: National Council of Teachers of English, 1967.

O'Rourke, J.P. *Toward a science of vocabulary development.* Paris: Mouton, 1974.

Pedersen, E. L. *Prospects for sentence-combining.* (Study prepared at Weber State, 1978.) (ERIC No.: ED 168 014.)

Perron, J.D. The impact of mode on written syntactic complexity. Athens, Ga.: Georgia University, 1976. (ERIC No.: ED 126 511.)

Powills, J.A., Bowers, R., and Conlan, G. Holistic essay scoring: An application of the model for the evaluation of writing ability and the measurement of growth in writing ability over time. Paper presented to annual meeting of the American Educational Research Association, April, 1979. (ERIC No.: ED 174 679.)

Reidy, E.F. *Fitchburg public schools skills achievement monitoring programs, writing.* Fitchburg, Mass: Fitchburg, Public Schools, 1980.

Rhode, M. New words in children's vocabulary. Paper presented at annual meeting of International Reading Association, May, 1977. (ERIC No.: ED 138 931.)

Rinsland, H. D. *A basic vocabulary of elementary school children*. New York, The Macmillan Company, 1945.

Rubin, R., and Buium, N. Language parameters in written composition of nine-year-old children. Paper presented to the annual meeting of the American Psychological Association, August, 1974, (ERIC No: 097 718.)

Ruddell, R.B. Language acquisition and the reading process. Paper presented at the annual meeting of the International Reading Association. Kansas City, 1969.

Schwartz, S., and Doehring, D.G. A developmental study of children's ability to acquire knowledge of spelling patterns. *Developmental Psychology*, 1977, *13*, 419-420.

Simon, D. P. Spelling—A task analysis, *instructional sciences*, 1976, *5*, 277-302.

Smith, C.B. First-grade composition as it relates to two methods of beginning reading in inner-city schools. Unpublished doctoral dissertation. Cleveland, Ohio: Case Western Reserve, 1967.

Smith, V. H. Composition rating scale, Bloomington, In.: School of Education, Indiana University: 1966. (ERIC No.: ED 091 721.)

Stewart, M. F., and Grobe, G.H. Syntactic maturity, mechanics of writing and teacher's quality ratings, *Research in the Teaching of English*, 1979, *13*, 207-215.

Suhor, C. Mass testing in composition: Is it worth doing badly? New Orleans, La.: New Orleans Public Schools, 1977. (ERIC No.: ED 147 807.)

Templeton, S. The circle game of English spelling: A reappraisal for teachers, *Language Arts*, 1979, *56*, 789-797.

Thompson, J.E. *Instructional intervention in the development of qualitative vocabulary: A further study*. (Doctoral dissertation, State University of New York at Buffalo, 1973), *Dissertation Abstracts International*. 1973. *34*, 1504A.

Thorndike, E.L. *Teacher's word book*. New York: Teachers College Press, 1921.

Thorndike, E.L., and Lorge, I. *The Teacher's word book of 30,000 words*. New York: Columbia University Press, 1944.

Thorndike, R.L. Reading as reasoning. *Reading Research Quarterly*, *9*, (1973-74), pp. 135-147.

VanDeWeghe, R. Research in written composition: Fifteen years of investigation, New Mexico State, University, 1978. (ERIC No.: ED 157 095.)

Veal, L.R. Syntactic measures and rated quality in the writing of children. *Studies in Language Education*, Report No., 8. Athens, Ga.: University of Georgia, 1974. (ERIC No.: ED 090 555.)

Walker, B. S. Vividness of imagery and spelling errors, *Perceptual and Motor Skills*, 1974, *39*, 823-825.

West, W. W. Writen composition. *Review of Educational Research*, 1967, *37*, 159-167.

Whitman, M.F. Welcome address. In A. Humes et. al., (eds.) *Moving between practice and research in writing. Proceedings of the NIE-FIPSE grantee workshop*. Los Alamitos: Calif: Southwest Regional Laboratory for Educational Research and Development, 1981.

Wing, A.M., and Baddeley, A.D. Spelling errors in handwriting: A corpus and a distributional analysis. In Frith, U. (Ed.) *Cognitive process in spelling*. New York: Academic Press, 1980.

Wirth, A. M., and Baddeley, A. D. Spelling errors in handwriting: A corpus and distributional analysis. In Uta Firth (Ed.) *Cognitive processes in spelling*. New York: Academic Press, 1980.

Yap, K.O. Vocabulary-building blocks of comprehensions? *Journal of Reading Behavior*, (1979), *11*, 50-59.

Yule, G.U. *The statistical study of vocabulary. Cambridge: University of Cambridge Press, 1944.*

Zutell, J. Some psycholinguistic perspectives on children's spelling, Language Arts, 1978, 55, 844-850.

500 Most Frequent

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
A	440	1717	3474	5946	6304	7598	8241	5879	39599
THE	636	1268	2241	3312	3353	4272	3839	3414	22335
AND	467	1159	2499	3498	3557	3773	3100	2443	20496
I	198	931	1230	2101	2609	2540	1875	1362	12846
TO	179	536	847	1803	2176	2446	1950	1834	11771
WAS	86	352	674	1047	1055	1376	1148	935	6673
MY	236	342	523	1081	1082	1001	1220	878	6363
OF	59	234	409	710	867	1213	1321	1042	5855
WE	122	205	439	1010	955	729	1253	801	5514
HE	210	364	781	790	883	1031	677	699	5435
IT	99	256	491	772	751	1106	1058	692	5225
THEY	85	167	468	981	830	1177	695	688	5091
WOULD	6	246	340	633	827	1136	786	544	4518
IS	129	288	422	601	721	738	877	573	4349
IN	71	214	296	315	572	824	948	736	3976
HAVE	25	120	263	504	645	751	721	646	3675
THAT	24	132	312	459	529	721	751	613	3541
FOR	27	106	339	421	515	653	684	563	3308
YOU	73	194	373	334	501	579	638	594	3286
SHE	43	147	328	551	425	617	578	498	3187
BE	11	111	198	283	568	637	676	560	3044
ON	48	149	313	386	499	566	602	418	2981
BUT	24	142	282	453	468	538	574	451	2932
WHEN	18	124	245	495	526	538	515	457	2918
ME	95	162	249	532	513	432	475	416	2874
LIKE	107	211	268	476	529	405	522	333	2851
THEN	41	179	345	484	447	469	476	320	2661
WERE	23	119	255	424	414	565	379	362	2541
ALL	17	137	182	423	379	502	365	417	2422
GO	85	151	243	337	389	409	414	289	2317
GET	52	137	172	309	381	453	398	408	2310
THERE	21	102	306	341	297	467	460	306	2300
WITH	70	116	192	323	329	438	421	358	2247
HAD	24	119	218	371	302	449	423	325	2231
ARE	30	85	191	397	374	405	436	299	2217
SO	17	96	211	367	359	447	352	313	2162
WENT	83	161	262	275	378	400	365	176	2100
UP	55	130	208	311	317	351	322	340	2034
AT	26	120	162	274	298	350	447	353	2030
SAID	36	132	254	391	429	362	232	176	2012
THEM	17	56	149	306	422	443	300	289	1982
IF	9	103	116	274	391	449	329	230	1901
HER	49	63	177	238	297	424	363	288	1899
ONE	9	86	233	256	276	403	353	266	1882
BECAUSE	29	153	181	312	317	283	346	235	1856
DO	27	111	149	258	333	326	342	281	1827
SCHOOL	61	188	146	179	264	329	336	318	1821
GOT	79	129	210	290	246	354	279	208	1795
HIS	49	131	267	207	291	325	203	241	1714
ABOUT	1	47	97	141	366	494	286	269	1701

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
DAY	56	118	166	297	248	351	250	178	1664
OUT	32	47	128	205	231	363	355	251	1612
HIM	29	115	184	211	248	329	234	203	1553
WILL	32	93	140	250	325	280	248	161	1529
NOT	21	109	163	207	225	265	264	223	1477
PEOPLE	10	87	158	150	195	257	290	260	1407
MAKE	13	28	129	177	260	266	218	177	1268
COULD	14	54	104	182	196	308	246	161	1265
OR	5	31	67	93	216	230	324	291	1257
CAN	95	87	80	135	179	194	303	173	1246
VERY	11	52	146	238	204	230	219	142	1242
PLAY	65	121	133	305	151	135	219	64	1193
SOME	16	69	104	168	193	216	241	176	1183
WHAT	15	54	89	136	172	297	193	227	1183
THIS	21	27	57	124	130	278	234	280	1151
TIME	8	66	110	173	177	208	206	195	1143
HOME	25	102	123	103	185	187	156	163	1134
GOING	42	92	139	195	163	178	189	126	1124
GOOD	16	64	108	144	198	187	234	154	1105
AS	4	25	64	98	136	236	261	268	1092
DOWN	61	116	133	137	146	183	149	146	1071
THEIR	16	24	103	149	147	237	157	210	1043
HOUSE	38	57	150	179	155	175	149	99	1002
BACK	6	102	103	137	143	178	176	153	998
CHARGE	1	74	77	158	170	222	167	96	965
CAME	32	65	108	185	152	166	137	111	956
FROM	3	36	75	116	152	201	173	189	945
FRIENDS	11	35	113	124	180	172	163	115	913
TOO	50	64	147	148	151	122	136	84	902
OTHER	4	35	74	98	145	178	217	129	880
AFTER	5	28	63	145	106	202	159	151	859
DON'T	6	34	86	110	171	141	153	146	847
OUR	2	23	38	121	119	108	249	158	818
NO	9	46	72	99	138	161	151	141	817
JUST	2	22	42	101	114	174	195	166	816
HAS	18	19	52	74	87	167	126	261	804
LOT	8	44	67	135	121	128	215	82	800
FUN	38	75	70	125	138	93	197	46	782
THINGS	1	49	54	117	102	144	187	127	781
BY	4	73	109	103	102	118	128	139	776
LITTLE	44	69	79	103	110	142	140	86	773
KNOW	6	39	71	105	122	143	129	139	754
WANT	14	39	54	120	163	117	139	100	746
SAW	40	60	109	141	86	132	79	92	739
FRIEND	14	33	106	85	145	118	138	92	731
DID	6	43	139	126	92	105	100	107	718
MORE	1	32	80	83	114	143	137	124	714
SEE	27	47	88	107	118	113	103	107	710
BIG	43	36	99	164	96	99	109	62	708
US	5	26	45	145	93	104	133	142	693

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
YOUR	4	34	79	68	105	110	126	166	692
EVERY	24	33	47	117	110	155	126	78	690
DIDN'T	2	23	68	84	100	160	137	115	689
TWO	9	22	49	108	107	166	139	71	671
DOG	70	46	90	125	87	101	82	58	659
HELP	20	54	59	102	95	156	80	91	657
MOTHER	15	47	99	113	106	100	87	85	652
AN	4	18	29	66	89	133	120	184	643
ALSO	0	10	113	62	71	133	145	107	641
AROUND	4	60	43	110	76	127	119	100	639
STARTED	5	12	50	83	92	110	163	120	635
NOW	5	19	53	101	101	126	116	100	621
THINK	5	18	48	58	89	97	159	132	606
COME	5	22	65	89	101	122	107	94	605
TAKE	6	28	31	108	99	113	101	116	602
NICE	21	37	94	104	130	92	69	53	600
FIRST	2	23	46	75	83	109	164	97	599
BEST	1	38	74	100	117	88	104	76	598
PUT	8	24	62	93	116	110	105	78	596
HOW	9	33	37	88	110	104	100	108	589
MAN	24	50	115	78	76	91	70	73	577
MOM	51	54	50	130	93	79	62	49	568
WHO	6	19	37	73	90	132	107	102	566
TELL	0	39	49	60	154	112	73	72	559
OVER	6	16	54	80	100	101	117	84	558
DAD	55	51	51	65	84	138	81	26	551
FAMILY	7	37	18	120	95	125	97	51	550
NAME	13	34	44	87	128	119	77	46	548
NEXT	3	18	57	95	91	96	104	68	532
NIGHT	7	15	47	108	81	112	99	57	526
MANY	0	11	13	56	47	81	153	163	524
LET	16	34	37	62	96	126	85	60	516
EAT	7	115	54	101	71	59	59	43	509
GIVE	7	56	23	55	116	94	83	69	503
TOLD	2	26	39	66	59	101	83	115	491
WORLD	6	13	99	69	84	84	67	66	488
RIGHT	5	17	46	86	68	99	79	78	478
AGAIN	1	24	148	69	56	64	56	48	466
TRY	2	16	24	50	67	106	113	75	453
WAY	1	15	32	68	65	99	89	84	453
WELL	1	17	28	76	77	81	92	80	452
LOVE	65	41	32	81	85	69	46	32	451
OFF	0	15	69	49	59	90	88	81	451
EVEN	3	58	37	41	48	78	102	75	442
THING	1	16	27	69	78	86	107	57	441
WORK	7	23	24	58	119	86	62	62	441
CLASS	0	17	8	63	102	143	41	66	440
WHERE	5	19	28	53	66	91	82	96	440
BOY	37	68	69	38	47	73	77	28	437
ANOTHER	2	16	45	47	75	76	102	69	432

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
RAN	26	34	50	66	48	98	51	54	427
ONLY	0	7	26	54	62	94	109	74	426
REALLY	6	11	22	50	64	78	106	87	424
FOOD	5	18	38	104	68	84	51	55	423
SOMETIMES	2	21	51	98	82	54	72	42	422
FOOTBALL	1	16	25	90	76	55	111	46	420
CALLED	4	15	26	52	70	136	65	51	419
FATHER	9	26	80	83	64	49	68	40	419
SOMETHING	2	16	39	54	56	79	79	91	416
TOOK	4	10	30	62	55	87	89	79	416
OLD	0	18	57	66	72	91	75	34	413
ONCE	2	38	80	64	68	66	59	34	411
NEW	12	11	40	39	66	90	81	70	409
CHILDREN	0	40	55	73	71	83	46	38	406
MUCH	8	32	43	58	47	80	82	55	405
CAR	16	24	57	62	47	111	49	28	394
INTO	10	27	38	51	56	84	50	77	393
MADE	5	5	50	78	57	70	75	53	393
RUN	17	12	26	52	77	86	61	61	392
YEARS	0	12	14	43	75	89	99	60	392
TEAM	2	3	9	55	75	38	162	46	390
KIDS	0	17	23	55	67	101	80	45	388
ALWAYS	1	19	17	73	57	74	78	67	386
AM	9	38	46	69	66	51	54	50	383
IT'S	3	15	33	33	77	78	89	48	376
WANTED	4	15	28	47	59	86	85	52	376
FOUND	3	16	46	48	44	73	84	60	374
BED	3	21	37	64	115	50	39	44	373
MONEY	3	9	36	47	85	79	51	61	371
WHY	3	20	20	44	94	81	60	49	371
NEVER	0	22	25	66	61	53	77	66	370
GIRL	17	54	42	44	68	80	37	26	368
AWAY	7	8	20	71	59	88	54	60	367
EACH	1	27	32	42	52	63	77	73	367
EVERYONE	1	23	24	49	61	66	74	67	365
ROOM	3	8	29	69	44	81	62	67	363
SISTER	9	19	24	62	83	52	58	53	360
ANY	1	8	25	40	63	76	73	73	359
TEACHER	1	30	43	43	72	54	66	50	359
THAT'S	2	21	27	50	69	75	73	41	358
FAVORITE	0	14	22	51	46	33	98	92	356
BROTHER	2	23	17	72	74	56	70	38	352
LONG	6	13	31	51	66	82	62	41	352
YEAR	0	14	23	25	56	69	109	54	350
GAME	3	16	30	41	79	42	107	31	349
MOST	0	11	17	57	54	50	86	73	348
CAT	55	21	52	87	40	39	26	21	341
HOMEWORK	17	122	6	15	30	28	27	85	330
GAMES	4	4	106	37	42	38	63	25	319
THOUGHT	0	7	30	56	39	52	74	61	319

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
SHOULD	0	5	12	21	36	51	69	124	318
BAD	1	30	27	48	47	42	78	42	315
CHRISTMAS	8	68	10	78	85	13	20	33	315
WATER	3	48	24	77	36	39	55	31	313
CLEAN	1	28	14	114	28	50	33	43	311
PARENTS	0	7	26	19	31	81	75	72	311
BEFORE	9	9	10	49	53	65	55	60	310
BETTER	0	6	28	47	45	63	61	58	308
I'M	1	13	33	38	71	40	66	40	302
LIVE	5	15	28	42	42	74	55	41	302
BUS	28	70	86	4	2	27	21	62	300
SAY	0	9	36	49	55	45	59	47	300
MORNING	4	17	32	39	41	53	51	62	299
STILL	0	11	25	42	56	60	57	47	298
HERE	14	19	23	34	51	57	55	44	297
LOOKED	3	22	31	44	28	59	42	67	296
WHILE	0	9	27	34	39	63	63	60	295
LEFT	1	6	25	62	47	51	56	46	294
STOP	1	29	55	23	36	60	41	47	292
AIR	8	25	31	20	25	27	132	21	289
CAN'T	15	18	14	106	37	29	32	37	288
THREE	23	9	23	40	52	68	36	37	288
HAPPY	5	29	45	42	57	45	35	28	286
EVERYBODY	3	18	18	48	51	47	63	36	284
EVERYTHING	2	3	26	36	56	43	54	61	281
UNTIL	3	26	22	41	53	44	47	44	280
ASKED	0	15	19	33	47	58	57	48	277
DIFFERENT	0	1	8	27	20	47	126	48	277
PLACE	2	13	19	44	46	54	40	59	277
SURE	1	5	20	36	37	55	52	70	276
NEED	1	24	38	29	52	42	38	50	274
GREAT	0	32	22	34	43	44	51	47	273
DOOR	7	9	41	35	34	55	43	48	272
LIFE	2	5	16	21	43	50	59	76	272
LOOK	9	16	39	39	32	58	40	38	271
SOMEONE	3	26	33	29	25	54	57	44	271
BALL	15	21	26	47	49	33	49	30	270
DAYS	1	11	13	80	30	66	31	36	268
WOULDN'T	0	10	12	13	53	68	68	43	267
STORY	0	19	26	25	36	71	22	62	261
FIND	3	11	22	29	38	58	46	53	260
FINALLY	4	7	16	44	29	63	53	41	257
TOGETHER	3	11	18	44	43	45	57	35	256
LIVED	2	24	37	56	45	53	19	16	252
ANYTHING	0	12	24	33	33	44	57	46	249
EVERY	6	11	23	32	47	60	41	28	248
BEEN	1	5	8	28	31	53	51	66	243
SWIMMING	1	5	16	44	37	33	53	54	243
KEEP	0	23	28	26	27	45	45	48	242
BUY	2	14	35	27	83	42	15	23	241

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
HEARD	0	10	24	35	37	57	41	37	241
THAN	0	8	26	31	42	44	53	37	241
GETTING	1	7	24	40	32	56	38	41	239
END	10	31	34	37	27	29	42	27	237
I'D	0	16	4	22	57	54	53	31	237
LAST	3	11	23	39	32	40	56	33	237
NAMED	5	12	22	40	58	55	31	14	237
TALK	1	2	7	44	42	51	48	42	237
COULDN'T	2	5	12	27	36	60	62	31	235
BOOKS	2	5	10	18	44	49	39	65	232
STAY	1	12	9	37	43	38	48	44	232
GIRLS	9	25	22	24	33	48	46	24	231
HARD	4	7	18	43	37	40	49	33	231
WHICH	0	8	9	13	26	50	58	66	230
YES	27	27	34	24	43	36	21	17	229
HIT	1	9	12	66	37	47	34	22	228
PERSON	0	14	28	26	46	26	55	33	228
ANIMALS	7	5	34	44	50	10	52	22	224
THROUGH	1	3	13	30	36	46	46	47	222
FELL	16	18	28	33	37	42	24	21	219
PLAYED	23	11	20	26	33	28	56	21	218
WISH	2	8	20	30	29	35	42	51	217
WATCH	1	14	24	48	36	32	31	30	216
BEING	1	6	8	15	42	30	52	61	215
KIND	0	5	18	36	36	33	50	37	215
WALKING	0	6	14	37	33	66	19	40	215
WHITE	11	11	15	41	44	51	27	15	215
IMPORTANT	0	3	2	94	80	9	14	11	213
HOPE	5	22	20	24	44	24	40	32	211
MEAN	4	67	20	16	40	23	21	17	208
WEEK	0	7	5	36	29	62	44	25	208
MIGHT	0	6	18	33	33	34	41	42	207
LOTS	4	24	33	34	43	25	31	12	206
KNEW	4	6	10	31	15	42	55	41	204
SCARED	8	23	32	33	27	35	29	17	204
BOYS	11	27	24	28	35	28	39	11	203
SOON	1	23	24	27	26	46	27	29	203
READ	0	14	5	16	36	47	32	52	202
LUNCH	2	17	12	18	23	41	42	45	200
STORE	7	11	14	29	46	45	40	18	200
WHOLE	5	8	11	33	36	39	36	31	199
BALLOONS	45	93	43	0	6	2	9	0	198
CALL	1	4	16	32	47	34	37	27	198
SEA	6	7	5	29	131	10	1	9	198
HORSES	1	2	10	16	103	14	43	8	197
BASEBALL	0	1	13	40	34	26	62	20	196
LATER	0	3	6	24	29	56	41	36	195
MEN	3	15	26	38	22	25	16	50	195
BEAR	1	10	9	109	34	19	7	5	194
REAL	3	4	13	22	33	38	39	42	194

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
START	0	3	14	28	25	55	38	31	194
ALMOST	0	7	18	49	33	38	25	23	193
PRETTY	0	9	31	28	32	31	38	24	193
HIGH	1	25	19	14	15	40	40	38	192
SAME	0	10	14	25	25	32	54	32	192
CARE	4	3	4	30	42	51	25	31	190
FEW	0	8	5	38	28	45	30	35	189
HORSE	0	7	18	17	66	25	36	20	189
DECIDED	0	3	9	21	27	43	44	41	188
HURT	2	11	28	40	32	26	19	30	188
BOOK	2	4	7	12	24	57	17	63	186
DOING	0	6	7	17	16	36	57	46	185
BLACK	18	5	14	20	15	77	23	12	184
RIDE	5	10	26	20	38	35	39	11	184
WALK	8	6	19	35	27	27	33	29	184
GAS	8	7	21	18	47	28	33	21	183
MR	0	14	24	27	45	38	20	15	183
TEACHERS	0	9	8	14	32	50	35	35	183
WASN'T	3	7	8	6	23	42	51	22	182
THESE	0	4	6	15	11	43	48	54	181
FISH	34	13	9	48	41	10	16	9	180
GOES	2	9	17	9	25	33	28	55	178
TREES	8	4	18	90	15	17	11	15	178
COMING	1	7	23	28	19	42	32	25	177
DREAM	0	3	16	9	8	44	38	58	176
GAVE	0	14	19	20	35	40	23	25	176
OUTSIDE	8	7	18	25	22	42	29	25	176
SLEEP	4	7	25	38	34	30	23	14	175
BOAT	7	48	8	58	19	9	18	7	174
GRADER	0	2	0	0	0	0	64	108	174
TREE	4	17	11	47	24	39	19	13	174
PRESIDENT	0	1	15	21	34	33	33	36	173
TV	9	14	29	36	22	23	21	18	172
PLAYING	5	9	24	28	26	26	29	24	171
MUST	2	0	3	15	11	30	24	85	170
SHOW	1	6	10	25	33	28	33	34	170
AFRAID	14	94	11	9	5	10	16	10	169
GETS	5	6	29	19	15	33	23	38	168
DOGS	2	13	17	28	23	36	34	14	167
STREET	1	9	22	28	19	32	26	29	166
CARS	3	6	20	33	19	31	35	18	165
LEARN	2	23	12	8	35	31	25	29	165
HE'S	3	6	20	22	24	42	21	26	164
TRIED	4	12	16	23	26	30	22	31	164
CANDY	3	5	16	37	24	41	26	11	163
FIGHT	0	7	14	31	35	27	20	29	163
LIKES	34	20	18	30	21	14	21	5	163
SNOW	76	0	10	24	17	14	18	4	163
DONE	2	3	8	17	22	27	37	45	161
USE	1	1	25	21	32	21	34	24	159

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
BABY	12	10	15	36	24	34	23	4	158
BASKETBALL	0	0	1	16	9	10	96	25	157
CIRCUS	16	57	59	5	9	1	10	0	157
FOUR	2	3	7	25	26	44	36	14	157
LOST	2	8	17	27	31	20	37	15	157
MAD	3	11	24	20	31	16	24	28	157
CLOTHES	0	7	11	28	30	27	27	26	156
GRADE	0	6	1	15	23	34	60	17	156
READY	3	9	6	17	33	32	20	34	154
TRIP	1	6	21	21	26	46	20	13	154
TURNED	3	9	14	29	9	26	27	37	154
WON	2	4	26	22	33	30	28	9	154
DOES	5	9	16	20	21	26	29	27	153
PROBABLY	0	1	10	9	23	18	64	28	153
DIED	0	4	25	41	25	27	9	2	152
FAST	9	6	14	21	31	40	17	14	152
OWN	0	1	8	21	34	23	37	28	152
WALKED	0	6	17	24	17	37	28	23	152
ASK	2	5	22	22	38	19	27	16	151
LAND	1	10	5	47	17	21	27	23	151
MAYBE	0	7	13	15	26	33	30	27	151
NOTHING	3	2	5	12	15	43	27	43	150
RUNNING	7	8	12	15	19	32	35	22	150
YOU'RE	0	7	12	23	27	25	33	23	150
MAY	5	1	10	11	20	31	22	49	149
BOTH	0	5	10	22	16	36	39	20	148
CITY	0	1	11	18	31	32	35	20	148
SHIP	0	2	10	54	16	48	5	13	148
FRIDAY	0	2	7	5	74	36	13	10	147
GROW	1	2	13	31	49	22	22	6	146
RED	12	8	19	26	17	29	19	16	146
BROKE	9	3	12	19	62	19	9	11	144
JOB	0	1	4	16	31	35	37	20	144
LOOKING	2	8	11	22	21	32	24	24	144
SMALL	0	5	11	16	18	39	31	24	144
TODAY	17	13	20	17	14	18	18	27	144
HAVING	1	4	3	14	19	26	34	42	143
JUMP	20	15	23	21	14	14	23	13	143
OK	2	13	22	26	30	27	16	7	143
PLANET	0	0	14	7	31	31	32	28	143
HOUR	0	0	6	20	30	38	35	13	142
I'LL	0	8	20	17	38	34	11	14	142
MYSELF	1	2	5	9	25	27	32	41	142
OH	1	9	19	2	35	24	19	14	142
EARTH	0	3	14	11	19	30	17	47	141
KILL	0	8	3	12	20	45	19	34	141
SUMMER	1	0	7	21	21	21	51	19	141
BEAUTIFUL	3	0	10	26	29	20	28	24	140
FUNNY	7	6	39	28	17	22	10	11	140
HAPPENED	6	7	14	9	10	35	30	28	139

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
PARK	13	4	9	25	28	15	34	10	138
UPON	2	39	20	23	19	19	9	7	138
EIGHTH	0	0	0	1	0	4	7	125	137
COMES	4	9	9	31	15	28	17	23	136
WAR	1	0	11	6	29	15	42	32	136
FEET	0	5	17	12	28	39	23	11	135
SET	0	3	7	13	14	33	21	44	135
WITHOUT	0	6	6	22	15	30	25	31	135
BRING	0	18	11	18	21	21	27	18	134
COUNTRY	0	2	5	18	21	19	39	30	134
ATE	4	7	24	32	17	15	21	13	133
CAUGHT	5	7	25	14	17	32	19	14	133
FIVE	0	9	10	24	19	29	30	12	133
CHANGE	1	6	1	15	13	22	45	29	132
RESPONSIBILITIES	0	0	0	1	0	1	3	127	132
THEY'RE	0	8	8	18	24	31	26	17	132
FIELD	2	4	14	8	28	29	36	10	131
LADY	0	2	25	35	13	20	16	20	131
MRS	0	13	9	19	31	24	18	17	131
TURN	3	2	38	16	11	23	17	21	131
ANIMAL	0	6	3	23	10	65	7	16	130
ENOUGH	0	3	10	15	18	19	33	32	130
TIMES	0	1	5	17	27	33	26	21	130
FREE	0	5	8	25	16	36	17	22	129
HEAD	6	5	12	20	15	27	20	24	129
SPORTS	0	1	9	32	18	19	31	19	129
BUILD	9	1	10	45	11	24	16	12	128
EXCEPT	0	17	2	15	22	26	23	23	128
BIKE	6	9	22	15	20	22	24	9	127
HALF	1	1	6	20	18	32	30	19	127
AMERICA	2	34	5	50	11	5	8	11	126
LEAVE	0	1	30	14	23	24	19	15	126
MOVE	0	10	13	21	21	30	23	8	126
POLICE	1	3	18	15	8	29	20	32	126
STATES	0	0	2	6	25	25	33	35	126
FEEL	0	1	5	17	16	26	30	30	125
STUFF	0	5	13	28	22	19	28	10	125
UNITED	0	3	4	4	26	18	34	36	125
MISS	1	14	14	27	11	12	33	12	124
SUDDENLY	5	2	10	15	16	32	11	33	124
TEACH	2	8	7	17	31	34	14	10	123
CATCH	1	4	14	32	31	10	16	13	121
FIRE	8	8	17	19	9	21	21	18	121
PARTY	0	4	8	27	23	27	19	13	121
DOESN'T	2	4	5	13	16	30	26	24	120
HAIR	0	4	9	23	19	25	16	24	120
PICK	2	7	11	20	17	13	32	18	120
REASON	0	6	1	6	19	31	40	17	120
SECOND	0	1	8	11	17	27	38	18	120
WINTER	51	9	12	18	7	7	13	3	120

APPENDIX A

WORD	GRADE								TOTAL
	1	2	3	4	5	6	7	8	
REST	0	2	8	10	26	29	27	17	119
KEPT	1	7	5	24	23	21	16	21	118
SIT	0	3	5	20	18	22	22	28	118
EYES	0	1	8	15	20	22	30	21	117
TOP	0	6	4	17	18	30	27	15	117
TROUBLE	0	3	2	19	9	26	32	26	117
FIX	2	0	0	18	7	15	60	14	116
FRONT	1	3	10	25	14	30	19	14	116
ELSE	1	2	6	16	12	17	31	30	115
HOT	2	5	33	12	9	21	20	13	115
MATH	1	28	12	13	15	24	18	4	115
SEEN	0	7	1	15	20	23	21	28	115
SHOT	1	6	10	24	22	26	10	16	115
DINNER	0	2	6	26	20	29	17	14	114
ITS	0	6	9	7	10	45	22	15	114
SICK	2	4	9	28	22	15	23	11	114
SINCE	0	3	2	16	19	22	30	22	114
SPACE	0	1	30	10	19	30	9	15	114
SPORT	0	1	0	32	16	2	46	17	114
SCHOOLS	1	4	6	11	19	24	34	14	113
THANKSGIVING	0	5	5	45	49	3	3	3	113
GROUND	3	9	14	16	11	23	19	17	112
SUCH	0	2	3	6	9	17	44	31	112
SUDDEN	1	3	9	21	14	30	14	20	112
TRYING	0	5	3	23	18	27	16	20	112
USED	0	2	3	13	15	29	33	17	112
FUTURE	0	17	15	2	22	21	17	16	110
MUSIC	0	4	6	12	10	27	20	31	110
PROBLEM	0	0	7	5	13	19	32	34	110
SEVENTH	0	0	0	0	0	2	105	3	110
WEEKS	0	3	6	19	16	25	26	14	109
WON'T	1	11	8	11	22	16	13	27	109
LIKED	2	10	13	13	13	21	19	17	108
LIVES	0	3	13	13	11	30	21	17	108
STOPPED	0	4	5	16	26	21	19	17	108
TALKING	1	3	10	19	10	35	14	16	108
THROW	15	5	6	26	11	20	11	14	108
WIN	1	3	8	20	17	19	23	17	108
WOKE	2	6	20	20	11	17	17	15	108
YARD	1	9	16	6	25	22	14	15	108
BELIEVE	0	14	3	12	17	17	23	21	107
GOD	6	4	6	29	25	9	12	16	107
LET'S	6	2	12	24	20	13	20	10	107
MOUSE	2	21	12	37	5	21	3	6	107
CUT	3	6	8	13	23	20	18	15	106
KILLED	0	10	7	8	17	28	14	22	106
MAKING	0	13	9	13	15	16	20	20	106
RIDING	9	7	21	8	18	7	30	6	106
RULES	1	3	3	13	23	25	20	18	106
BECOME	0	0	1	16	18	26	21	23	105

Common misspellings

APPENDIX B

WORD	MISSPELLING	GRADE								TOTAL
		1	2	3	4	5	6	7	8	
A LOT	ALOT	2	5	29	42	28	44	39	33	222
A WHILE	AWHILE	0	0	0	0	0	8	3	4	15
AGAIN	AGIAN	0	0	2	6	3	4	0	2	17
AGAIN	AGIN	0	4	1	2	1	3	0	1	12
ALL RIGHT	ALRIGHT	0	0	0	2	1	7	2	4	16
ALSO	ALLSO	0	1	6	2	1	0	0	0	10
ALWAYS	ALLWAYS	0	1	0	1	3	7	2	0	14
AM	AN	0	0	4	5	4	1	0	0	14
AND	A	3	2	4	5	6	7	3	2	32
AND	AN	5	10	7	13	29	6	4	0	74
AND	IN	3	12	0	5	15	4	2	1	42
ANOTHER	A NOTHER	0	0	2	5	3	5	4	2	21
ANY MORE	ANYMORE	0	1	0	2	2	4	2	5	16
ANYONE	ANY ONE	0	0	0	1	1	5	4	2	13
ANYTHING	ANY THING	0	0	3	0	2	2	3	2	12
ANYWAY	ANY WAY	0	0	1	0	3	3	2	1	10
AROUND	AROND	1	3	5	0	0	1	0	0	10
BALLOONS	BALLONS	0	11	0	0	0	0	0	0	11
BASKETBALL	BASKET BALL	0	0	0	0	1	1	16	1	19
BECAUSE	BECASE	0	1	5	7	8	3	0	0	24
BECAUSE	BECAUS	0	2	4	4	4	4	4	2	24
BECAUSE	BECOUSE	0	1	3	1	3	0	3	6	17
BECAUSE	BECUSE	0	8	6	4	5	0	0	1	24
BECAUSE	CAUSE	0	3	1	14	2	1	12	13	46
BEFORE	BEFOR	1	0	3	1	5	2	0	1	13
BELIEVE	BELIVE	0	1	0	1	4	1	2	1	10
BIGGER	BIGER	0	0	8	0	2	1	1	0	12
BLACK HOLE	BLACKHOLE	0	0	0	0	0	10	0	0	10
BUY	BY	0	4	3	5	7	3	1	0	23
CAME	CANE	0	0	1	1	1	6	1	0	10
CAN'T	CANT	0	1	3	1	0	3	4	4	16
CANNOT	CAN NOT	2	1	1	2	2	3	0	0	11
CAPTAIN	CAPTIAN	0	0	0	4	0	7	1	0	12
CHARGE	CARGE	0	0	0	7	4	0	0	0	11
CHARGE	CHARG	0	1	0	6	1	4	0	0	12
CHOIR	CHIOR	0	3	0	0	0	0	5	2	10
CHRISTMAS	CHRISMAS	0	2	0	6	0	1	1	1	11
CLASSROOM	CLASS ROOM	0	0	1	0	2	8	2	1	14
CLOTHES	CLOSE	0	1	0	5	2	3	1	0	12
CLOTHES	CLOTHS	0	1	0	4	2	3	2	0	12
DIDN'T	DID'NT	0	0	3	3	0	6	1	0	13
DIDN'T	DIDNT	0	1	1	10	2	7	2	3	26
DIFFERENT	DIFFRENT	0	0	2	4	2	3	6	0	17
DOGS	DOG'S	0	0	1	4	0	6	0	0	11
DON'T	DONT	0	2	16	7	5	9	12	16	67
DOWN	DAN	9	4	0	0	0	0	0	0	13
DOWNSTAIRS	DOWN STAIR	0	0	0	1	4	1	2	3	11
EIGHTH	EIGTH	0	0	0	0	0	0	3	14	17
ETC	ECT	0	0	1	0	1	1	6	7	16

APPENDIX B

WORD	MISSPELLING	GRADE								TOTAL
		1	2	3	4	5	6	7	8	
EVERY	EVEY	0	0	2	4	3	0	1	2	12
EVERY DAY	EVERYDAY	0	0	0	4	4	6	5	3	22
EVERYBODY	EVERY BODY	0	4	2	10	6	6	5	0	33
EVERYONE	EVERY ONE	0	0	7	10	7	10	3	5	42
EVERYTHING	EVERY THING	0	0	1	3	4	4	2	3	17
EVERYWHERE	EVERY WHERE	0	1	4	4	1	0	2	0	12
FAIR	FARE	0	2	0	2	1	5	1	0	11
FAMILY	FAMLY	6	2	0	2	0	0	0	0	10
FELL	FEL	4	4	2	0	0	0	0	0	10
FIELD	FEILD	0	0	4	2	4	9	0	1	20
FINALLY	FINALY	0	0	3	10	0	8	0	0	21
FINALLY	FINELY	0	1	0	8	0	3	0	0	12
FIRST	FRIST	0	0	2	4	2	3	0	0	11
FOOTBALL	FOOT BALL	0	2	12	2	0	0	4	0	20
FOUND	FOND	0	3	10	1	0	4	0	0	18
FRIEND	FREIND	0	2	9	12	12	9	7	3	54
FRIEND'S	FRIENDS	0	0	1	4	7	4	0	1	17
FRIENDS	FREINDS	0	0	8	13	22	8	3	2	56
FRIENDS	FRENDS	0	3	1	3	3	0	0	0	10
FRIENDS	FRIEND'S	0	4	5	1	0	0	2	2	14
FRIENDS	FRINDS	0	1	0	8	2	0	4	0	15
GET	GIT	5	0	4	0	0	1	0	0	10
GETTING	GETING	0	0	3	2	2	2	0	1	10
HAPPENED	HAPPEND	0	1	1	2	4	6	1	0	15
HAPPILY	HAPPLY	1	2	6	8	2	0	0	1	20
HAVE	HAV	4	7	0	0	2	0	0	0	13
HE'S	HES	0	0	2	0	6	3	0	2	13
HEARD	HERD	0	4	9	2	3	7	1	9	35
HER	HE	0	0	0	0	7	2	1	3	13
HERE	HEAR	0	0	2	4	0	2	1	2	11
HERO	HEROE	0	0	6	1	9	0	0	0	16
HEROES	HEROS	0	0	8	0	2	4	0	2	16
HIM	HEM	5	4	4	0	0	11	0	0	24
HIM	HIN	0	2	9	4	0	4	0	0	19
HIS	HE'S	0	0	2	2	4	0	0	2	10
HIS	HES	9	11	8	2	0	0	0	0	30
HIS	IS	0	4	5	2	0	0	0	2	13
HOMEWORK	HOME WORK	0	0	1	4	4	2	3	2	16
HURT	HERT	0	2	3	3	2	0	0	4	14
I'D	ID	0	2	0	4	2	12	1	1	22
I'LL	ILL	0	0	4	4	10	8	2	3	31
I'M	IM	0	3	9	3	10	5	6	10	46
IN	I	0	6	0	2	1	0	2	0	11
IN CHARGE	INCHARGE	0	0	0	2	4	16	10	0	32
INDIANS	INDENS	0	0	0	6	4	0	0	0	10
INSIDE	IN SIDE	0	2	2	1	6	3	0	0	14
INSTEAD	INSTED	0	0	5	0	2	3	0	0	10
INTO	IN TO	4	3	2	11	12	12	4	4	52
IT'S	ITS	0	6	10	9	16	22	30	23	116
ITS	IT'S	0	7	3	2	4	10	5	2	33

APPENDIX B

WORD	MISSPELLING	GRADE								TOTAL
		1	2	3	4	5	6	7	8	
KNEW	NEW	0	1	4	5	1	1	1	1	14
KNOW	NO	5	12	4	13	10	2	1	1	48
KNOW	NOW	0	1	5	4	11	2	3	1	27
LET'S	LETS	1	0	7	5	5	2	9	4	33
LITTLE	LITTEL	0	1	6	1	2	0	0	0	10
LYING	LAYING	0	0	0	0	3	5	0	5	13
MAKE	MACK	0	2	9	1	2	1	0	0	15
MAYFLOWER	MAY FLOWER	0	1	0	11	0	0	0	0	12
ME	MY	7	0	7	1	2	0	0	0	17
MIGHT	MITE	0	1	4	5	0	1	0	0	11
MILLIMETER	MILIMETER	0	0	0	0	0	10	0	0	10
MOTHER'S	MOTHERS	0	2	0	4	1	2	1	4	14
MYSELF	MY SELF	0	0	3	0	5	1	4	2	15
NEW	NOW	7	0	0	3	0	0	0	0	10
NO	KNOW	0	0	0	6	0	3	2	1	12
NOW	KNOW	0	0	3	1	3	2	1	3	13
OFF	OF	0	1	2	3	6	11	4	6	33
ONCE	ONES	0	1	6	1	1	0	1	0	10
ONE	ON	1	1	1	4	1	2	1	3	14
OR	ORE	0	1	1	1	4	0	1	10	18
OUR	ARE	0	4	10	25	16	8	7	3	73
OUTSIDE	OUT SIDE	1	4	5	3	4	4	1	1	23
PASSED	PAST	0	1	2	6	5	3	1	2	20
PEOPLE	PEAPLE	0	1	4	3	3	0	0	0	11
PEOPLE	PEOPL	0	0	0	2	7	1	1	0	11
PEOPLE'S	PEOPLES	0	1	1	4	0	2	2	2	12
PILGRIMS	PILGRAMS	2	2	0	6	0	0	0	0	10
PLAY	PALY	5	0	1	7	0	0	0	0	13
POPPED	POPED	0	7	4	0	0	1	0	0	12
PROBABLY	PROBLY	0	0	2	0	6	4	5	0	17
REALLY	REALY	0	2	8	8	5	3	4	3	33
SAID	SED	12	13	10	0	0	0	1	0	36
SAID	SIAD	0	2	5	6	5	5	0	0	23
SAID	SIDE	0	6	2	0	4	0	0	0	12
SATURDAY	SATERDAY	0	1	1	1	7	1	1	0	12
SCHOOL	SHCOOL	5	0	0	2	0	9	1	1	18
SCHOOL	SHOOL	0	1	0	2	3	4	2	1	13
SHE'S	SHES	0	2	11	0	3	1	0	1	18
SOME	SOM	4	1	1	2	2	0	0	1	11
SOME	SUM	6	7	3	2	0	0	1	0	19
SOMEONE	SOME ONE	0	1	1	0	0	5	1	5	13
SOMETHING	SOMTHING	0	2	6	4	4	6	1	0	23
SOMETIMES	SOME TIMES	0	0	4	5	7	0	2	1	19
SOMETIMES	SOMTIMES	0	0	1	3	6	1	1	0	12
STARTED	STARED	0	0	3	4	2	1	1	1	12
STOPPED	STOPED	0	1	1	4	6	2	2	2	18
STRETCH	STRECH	0	9	0	1	0	2	0	0	12
SURE	SHARE	0	0	0	0	0	0	0	11	11
SWIMMING	SWIMING	0	2	7	11	3	1	2	1	27
TEACHER	TECHER	0	4	4	1	2	0	1	0	12
THAN	THEN	0	4	17	1	5	0	3	2	32

APPENDIX B

WORD	MISSPELLING	GRADE								TOTAL
		1	2	3	4	5	6	7	8	
THAT'S	THATS	0	6	11	8	20	21	20	17	103
THEIR	THER	0	3	2	2	1	0	4	1	13
THEIR	THERE	0	5	4	23	14	16	15	14	91
THEIR	THIER	0	0	0	3	3	4	13	5	28
THEM	THEN	0	0	3	7	2	5	0	0	17
THEN	THAN	0	1	0	0	4	7	4	2	18
THEN	THE	2	0	1	5	1	3	0	2	14
THERE	THEIR	0	4	6	10	12	4	7	5	48
THERE	THER	1	9	8	5	0	2	2	1	28
THEY	TAY	7	0	5	0	0	0	0	0	12
THEY	THAE	9	1	0	0	0	0	0	0	10
THEY	THAY	10	16	16	14	15	1	0	0	72
THEY	THE	2	0	3	7	4	6	8	10	40
THEY'RE	THEIR	0	0	1	0	2	1	5	4	13
THEY'RE	THERE	0	1	3	1	2	7	3	3	20
TO	TOO	3	2	2	9	6	4	16	8	50
TODAY	TO DAY	8	0	0	0	2	0	1	0	11
TOO	TO	15	39	75	47	50	22	42	28	340
TURKEY	TURKY	0	0	0	4	10	0	0	0	14
TWO	TO	2	2	1	6	7	0	2	1	21
TWO	TOW	0	5	1	15	3	0	0	0	24
UNTIL	UNTILL	0	0	1	1	5	7	6	1	21
UPON	APON	0	1	2	2	6	1	0	1	13
WE'RE	WERE	0	0	3	11	4	2	4	8	32
WENT	WANT	3	5	2	0	0	0	0	0	10
WENT	WET	11	2	0	0	0	0	0	0	13
WENT	WHENT	0	0	5	3	0	3	0	2	13
WERE	WHERE	0	14	17	6	5	6	3	5	56
WHAT'S	WHATS	0	2	2	5	4	6	2	1	22
WHATEVER	WHAT EVER	0	0	2	3	5	0	3	6	19
WHEN	WEN	1	0	9	4	6	2	0	0	22
WHEN	WIN	0	14	4	0	1	2	0	1	22
WHENEVER	WHEN EVER	0	0	1	2	3	2	4	0	12
WHERE	WERE	0	3	4	8	10	16	9	7	57
WHILE	WILE	0	2	5	1	2	2	0	0	12
WHOLE	HOLE	0	2	4	8	5	3	0	0	22
WITH	WHITH	0	4	5	2	3	2	2	0	18
WITH	WIHT	5	2	12	0	0	0	0	0	19
WON'T	WONT	0	1	2	0	5	1	2	2	13
WORLD	WOLD	0	0	5	1	2	2	0	0	10
WOULD	WODE	0	6	2	1	1	1	0	0	11
WOULD	WOLD	0	8	8	7	7	1	2	1	34
WOULD	WOOD	2	15	4	15	10	17	2	0	65
WOULD	WOUD	0	4	6	3	14	2	1	0	30
WOULD	WUD	3	6	1	0	0	0	0	0	10
WOULDN'T	WOULDNT	0	0	1	0	3	4	6	3	17
YOU'RE	YOUR	0	2	8	14	9	8	16	13	70

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 21,1991